

UNIVERSITÉ DU QUÉBEC À MONTRÉAL

CARACTÉRISATION DES LEUCÉMIES INDUITES PAR LE RÉTROVIRUS
GRAFFI

THÈSE
PRÉSENTÉE
COMME EXIGENCE PARTIELLE
DU DOCTORAT EN BIOLOGIE

PAR
VÉRONIQUE VOISIN

Septembre 2008

UNIVERSITÉ DU QUÉBEC À MONTRÉAL
Service des bibliothèques

Avertissement

La diffusion de cette thèse se fait dans le respect des droits de son auteur, qui a signé le formulaire *Autorisation de reproduire et de diffuser un travail de recherche de cycles supérieurs* (SDU-522 – Rév.01-2006). Cette autorisation stipule que «conformément à l'article 11 du Règlement no 8 des études de cycles supérieurs, [l'auteur] concède à l'Université du Québec à Montréal une licence non exclusive d'utilisation et de publication de la totalité ou d'une partie importante de [son] travail de recherche pour des fins pédagogiques et non commerciales. Plus précisément, [l'auteur] autorise l'Université du Québec à Montréal à reproduire, diffuser, prêter, distribuer ou vendre des copies de [son] travail de recherche à des fins non commerciales sur quelque support que ce soit, y compris l'Internet. Cette licence et cette autorisation n'entraînent pas une renonciation de [la] part [de l'auteur] à [ses] droits moraux ni à [ses] droits de propriété intellectuelle. Sauf entente contraire, [l'auteur] conserve la liberté de diffuser et de commercialiser ou non ce travail dont [il] possède un exemplaire.»

REMERCIEMENTS

Je tiens à remercier tout particulièrement le Dr. Eric Rassart de m'avoir accueilli dans son laboratoire de biologie moléculaire du département des Sciences Biologiques de l'Université du Québec à Montréal en Juin 2001. Je le remercie premièrement de m'avoir donné l'occasion de faire face à un projet complexe et donc très stimulant. Je le remercie de m'avoir accordé beaucoup de confiance tout au long du projet tout en m'appuyant dans quelques moments délicats. Merci également de m'avoir donné l'occasion de compléter ma formation en me permettant d'encadrer des étudiants, de pratiquer l'enseignement et de présenter mes travaux au cours de congrès internationaux.

Je remercie également tous les professeurs avec qui j'ai pu communiquer et qui m'ont toujours apporté une aide, les membres du jury qui vont lire la présente thèse, ainsi que les Drs. L. Poliquin et T. Hoang qui ont suivi mon parcours en tant que membres du comité de thèse.

Je remercie toutes les personnes du laboratoire présentes au cours de mes années de doctorat. Je remercie tout particulièrement Diana Paulina Ospina Salazar et Cyndia Charfi pour les périodes pendant lesquelles nous avons travaillé conjointement sur le projet. Je remercie Sonia Do Carmo et Philippe Legault, qui m'ont aidé en tant qu'assistants de recherche du laboratoire. Je tiens également à remercier Séverine Landais et Fée-ann C. McNabb pour les échanges tant sur le point scientifique qu'un peu plus amical.

Finalement, un gros merci à toute ma famille et mes proches pour leur énorme soutien.

TABLES DES MATIÈRES

Remerciements	ii
Table des matières	iii
Liste des figures	vi
Liste des tableaux	x
Abréviations	xii
Résumé	xiii
 CHAPITRE I: ÉTAT DES CONNAISSANCES	 1
1.1 Hématopoïèse et Leucémie	1
1.1 Hématopoïèse.....	1
1.1.2 Leucémie.....	30
1.2 Les rétrovirus murins de la leucémie	47
1.2.1 Découverte des rétrovirus murins de la leucémie	47
1.2.2 Structure des MuLVs	49
1.2.3 Cycle de réplication rétrovirale.....	51
1.2.4 Induction de la leucémie	52
1.2.5 Le rétrovirus Graffi	56
 CHAPITRE II: CARACTÉRISATION DES LEUCÉMIES INDUITES PAR LE RÉTROVIRUS GRAFFI.....	 65
2.1 Avant-propos.....	65
2.2 Article: 'Novel insights into the pathogenesis of the Graffi Murine Leukemia Retrovirus'	66
2.2.1 Abstract.....	66
2.2.2 Introduction.....	67
2.2.3 Materials and Methods.....	68
2.2.4 Results.....	70
2.2.5 Discussion	78
2.2.6 Acknowledgments.....	84
2.3 Résultats supplémentaires.....	85

2.3.1 Schématisation plus visuelle des différentes leucémies induites par le rétrovirus Graffi.....	85
2.3.2 Clonage et séquençage du enhancer dans différentes tumeurs.....	85
2.2.3 Résultats complémentaires sur la caractérisation des leucémies.....	87
2.3.4 Conclusion	89
2.4 Figures et tableaux du chapitre II.....	90
 CHAPITRE III: ANALYSES PHYLOGÉNÉTIQUES	101
3.1 Avant-propos.....	101
3.1.1 Introduction.....	101
3.1.2 Principes de base de la phylogénie moléculaire.....	102
3.2 Article: ' <i>Complete genome sequences of the two viral variants of the Graffi MuLV: Phylogenetic relationship with other murine leukemia retroviruses</i> '	104
3.2.1 Abstract.....	104
3.2.2 Introduction.....	105
3.2.3 Material and methods.....	106
3.2.4 Results.....	108
3.2.5 Discussion	117
3.2.6 Acknowledgments.....	122
3.3 Résultats supplémentaires: Distribution géographique des MuLVs	123
3.4 Figures et tableaux du chapitre III	125
 CHAPITRE IV: PROFIL D'EXPRESSION GÉNIQUE DES LEUCÉMIES INDUITES PAR LE RÉTROVIRUS GRAFFI.....	133
4.1 Introduction.....	133
4.1.1 Stratégie générale	133
4.1.2 Présentation d'une micropuce	134
4.1.3 Tri des leucémies	138
4.2 Phénotypes des leucémies triées et envoyées à l'analyse par micropuces	138
4.3 Résultats Partie I: validation de l'expérience.....	140

4.3.1	Présentation des données brutes.....	140
4.3.2	Présentation des données normalisées	140
4.3.3	Vérification de la validité de l'expérience	144
4.4	Résultats Partie II: Profil général des leucémies.....	146
4.4.1	Analyse différentielle.....	146
4.4.2	Sélection de 3000 gènes.....	148
4.4.3	Clustering hiérarchique, analyse SOMs et profils des leucémies	148
4.5	Résultats partie III: gènes spécifiques de chaque type de leucémies	154
4.5.1	Méthodologie	154
4.5.2	Analyse ciblée sur les leucémies non-lymphoïdes.....	156
4.5.3	Analyse ciblée sur les leucémies lymphoïdes	205
4.6	Résultats partie IV: Validation par RT-PCR.....	221
4.6.1	Stratégie générale.....	221
4.6.2	Matériel et méthodes.....	222
4.6.3	Résultats des expériences de RT-PCR	224
4.7	Recherche de sites d'intégrations.....	251
4.7.1	Matériel et méthode	251
4.7.2	Résultats.....	253
4.7.3	Vérification des sites d'intégrations 5E1 et 6D2	256
4.8	Discussion et perspectives	261
4.8.1	Résumé.....	261
4.8.2	Perspectives.....	262
4.8.3	Conclusion	271
4.9	Remerciements.....	272
	RÉFÉRENCES.....	273

LISTE DES FIGURES

Figure	Page
1.1	Schéma général de l'hématopoïèse.....5
1.2	Organisation du microenvironnement de la moelle osseuse.....7
1.3	Antagonisme entre GATA-1 et PU.1.....10
1.4	Le développement des cellules B.12
1.5	Le développement des cellules T.....15
1.6	L'érythropoïèse.....21
1.7	La mégacaryopoïèse.22
1.8	Concept du continuum et de l'état 'cellule souche'28
1.9	Comparaison de l'état de différenciation.....37
1.10	Cluster hiérarchique de 327 échantillons de leucémies ALL.46
1.11	Organisation des virions et des génomes viraux.....50
1.12	Cycle de réplication rétrovirale.....52
1.13	Mécanismes de la mutagenèse insertionnelle.54
1.14	Arnold Graffi et le virus Graffi.....57
1.15	Cartes de restriction de GV-1.4, GV-1.2 et GV-2.5.60
1.16	Activité de transcription de la région U3 de plusieurs virus.....62
1.17	Intégrations rétrovirales trouvées dans des tumeurs induites par GV-1.4 64
2.1	Survival analysis of newborn BALB/c, NFS, and FVB/n mice ... injected with the molecular clones GV-1.4 90

2.2	Flow cytometry analyses of several types of leukemias.....	91
2.3	Southern blot analysis of TCR β (A) and immunoglobulin..... heavy chain (B) gene rearrangement.	92
2.4	Northern blot analysis of some specific genes of the myeloid, erythroid, and megakaryocytic lineages.	93
2.5	Agarose gel analysis of proviral enhancer PCR amplification in mature tumors.....	94
2.6	Répartition des différents types de leucémies induites..... par le rétrovirus Graffi... ..	97
2.7	Multimérisation du enhancer, clonage et séquençage.....	98
2.8	Caractérisation complémentaire des leucémies	99
2.9	Coloration de lames	100
3.1	Phylogenetic analysis of several MuLV genomes.....	125
3.2	Phylogenetic analysis strategy and detection of hypervariable regions.....	126
3.3	Phylogenetic analysis strategy and detection of hypervariable regions.....	127
3.4	Sequential phylogenetic analysis of the Pol gene.	128
3.5	Region Binding Domain of SU.....	129
3.6	Phylogenetic analysis and multiple alignment of the LTR..... region.	130
3.7	Distribution géographique des MuLVs et des souches..... de souris.....	132
4.1	Présentation d'une micropuce Affymetrix.....	135
4.2	Schématisation de la procédure d'Affymetrix	137
4.3	Images des données brutes des micropuces	141
4.4	Graphiques de distribution des intensités de chaque micropuce ..	142
4.5	Graphiques PCA	145
4.6	Visualisation des clusters hiérarchiques	150

4.7	Profil général des leucémies	152
4.8	Schéma du complexe de l'intégrine $\alpha\text{IIb}\beta 3$ et de la biosynthèse du hème.....	162
4.9	Récepteurs et signalisation impliqués dans la phagocytose d'un microbe	179
4.10	Phagocytose par les mégacaryocytes et les plaquettes	180
4.11	Schéma des gènes impliqués dans la réponse immunitaire exprimés par les cellules mégacaryoblastiques	181
4.12	Gènes impliqués dans la voie de signalisation du TCR.....	208
4.13	Différences entre T1, T2 et T3	211
4.14	Résultats de RT-PCR pour Gulp1	226
4.15	Résultats de RT-PCR pour Fkbp9	226
4.16	Résultats de RT-PCR pour Gucyl a3	228
4.17	Résultats de RT-PCR pour Ncf2.....	228
4.18	Résultats de RT-PCR pour Kit	230
4.19	Résultats de RT-PCR pour Gata2	230
4.20	Résultats de RT-PCR pour Gata1	232
4.21	Résultats de RT-PCR pour Snca.....	232
4.22	Résultats de RT-PCR pour Slamf1	234
4.23	Résultats de RT-PCR pour Ltpb2	234
4.24	Résultats de RT-PCR pour Cda	238
4.25	Résultats de RT-PCR pour Btbd14a.....	238
4.26	Résultats de RT-PCR pour Rabgef1	240
4.27	Expression des gènes spécifiques des leucémies mégacaryo-blastiques dans des échantillons non leucémiques	241
4.28	Expression des gènes spécifiques des leucémies érythroïdes- dans des échantillons non leucémiques	242
4.29	Résultats de la RT-PCR pour Fbln1	244

4.30	Résultats de la RT-PCR pour Bmp7	244
4.31	Résultats de la RT-PCR pour Rgs3	246
4.32	Résultats de la RT-PCR pour Hdgfrp3	246
4.33	Résultats de la RT-PCR pour Etv5	248
4.34	Résultats de la RT-PCR pour Jakmip1	248
4.35	RT-PCR pour de 2 gènes spécifiques des leucémies B	250
4.36	Recherche de sites d'intégrations	252
4.37	Sites d'intégrations 5E1 et 6D2	259
4.38	RT-PCR de Kit et GATA-2 dans Mk1, Mk2 et Mk3	260

LISTE DES TABLEAUX

Tableau	Page
1.1	Classification des néoplasmes humains lymphoïdes33
1.2	Classification des leucémies humaines myéloïdes aiguës33
1.3	Désordres hématopoïétiques lymphoïdes chez la souris.....35
1.4	Désordres hématopoïétiques non-lymphoïdes chez la souris35
2.1	Gross pathology and flow cytometry results of the leukemic..... mice 95
2.2	Global analysis of the T- and B- leukemias.....96
3.1	List of the MuLVs used in this study.....131
4.1	Leucémies hybridées sur micropuces139
4.2	Explication du calcul des ratios147
4.3	Groupe(s) identifiées par analyse SOM153
4.4	Vérification des phénotypes E et Mk.....157
4.5	Facteurs de transcription spécifiques des E et Mk.....160
4.6	Itga2b et gènes reliés163
4.7	Synthèse du hème par les érythrocytes165
4.8	Evi.1.....170
4.9	Gènes mégacaryocytaires sélectionnés pour l'analyse RT-PCR ..182
4.10	Expression de Lrp1183
4.11	Expression de Gucy1b3185
4.12	La NADPH oxydase186

4.13	Hemt	193
4.14	Gènes érythrocytaires sélectionnés pour la validation RT-PCR...	197
4.15	Expression des Ltbp et TGFb.....	198
4.16	Expression de Sh2d1a.....	200
4.17	Expression de Rabgef1 et Rab22	202
4.18	Phénotypes des leucémies T	206
4.19	Vérification des phénotypes des leucémies B.....	212
4.20	Présentation de gènes d'intérêt pour les leucémies B	215
4.21	Exemples de gènes spécifiques des leucémies T et B	216
4.22	Ratios d'expression de Lmo4	219
4.23	Phénotypes des gènes testés en RT-PCR.....	222
4.24	Gènes non-lymphoïdes sélectionnés.....	224
4.25	Résumé de l'expression des gènes non-lymphoïdes	237
4.26	Gènes lymphoïdes sélectionnés	243
4.27	Intégrations rétrovirales dans Mk1	253
4.28	Intégrations rétrovirales dans Mk2	254
4.29	Intégrations rétrovirales dans Mk3	254
4.30	Sites d'insertion et résultats des micopuces.....	255

LISTE DES ABRÉVIATIONS

ALL	leucémies aiguës lymphoïdes
AMKL	leucémies mégacaryoblastiques aiguës
AML	leucémies aiguës myéloïdes
APL	leucémies aiguës promyélocytiques
BCR	récepteur des cellules B
CLL	leucémies chroniques lymphoïdes
CLP	progéniteur commun lymphoïde
CML	leucémies myéloïdes chroniques
CMP	progéniteur commun myéloïde
DN	double-négatif
EPO	érythropoïétine
HSC	cellule souche hématopoïétique
LTR	<i>long terminal repeat</i>
MEP	progéniteur bipotent érythro-mégacaryocytaire
MPO	myéloperoxidase
MuLV	rétrovirus murins de la leucémie
NK	<i>natural killer</i>
OMS	organisation mondiale de la santé
RAR	récepteur α de l'acide rétinoïque
RBD	<i>region binding domain</i>
SAM	<i>significance analysis of microarrays</i>
SOM	<i>self organization map</i>
TCR	récepteur des cellules T
TPO	thrombopoïétine

RÉSUMÉ

Le rétrovirus Graffi induit des leucémies chez la souris et permet donc l'étude de phénomènes impliqués dans la leucémie ou plus généralement dans le cancer. Le but du présent projet a été de caractériser le rétrovirus murin Graffi et les leucémies induites afin d'apporter des éléments nouveaux aidant à la compréhension de la leucémie.

La leucémie est une conséquence d'un dysfonctionnement de l'hématopoïèse. Les cellules leucémiques dérivent de progéniteurs hématopoïétiques qui ont subi des mutations et l'expression de plusieurs proto-oncogènes est dérégulée. Ainsi ces cellules anormales ne peuvent pas terminer le processus de différenciation, prolifèrent de façon anarchique, échappent à l'apoptose et envahissent l'organisme. Bien que ne formant pas de tumeurs solides, les cellules leucémiques ont beaucoup de caractéristiques en commun avec d'autres types de cellules cancéreuses.

Le rétrovirus murin Graffi a été isolé en 1954 par Arnold Graffi, qui l'a caractérisé comme un rétrovirus induisant des leucémies myéloïdes. Son équipe et lui-même ont cependant été confrontés au phénomène qu'ils ont appelé 'diversification hématologique'. Récemment, 2 clones moléculaires ont été dérivés dans le laboratoire d'Eric Rassart à partir de l'extrait original du rétrovirus Graffi et qui correspondent à 2 génomes non défectifs, nommés GV-1.2 et GV-1.4. Les leucémies induites par ces 2 variants avaient été classifiées comme myéloïdes, ceci basé sur l'observation morphologique de frottis sanguins. Une nouvelle étude de caractérisation des leucémies induites par ces 2 variants dans 3 souches de souris a montré que le rétrovirus Graffi est capable d'induire des types de leucémies très variés: lymphoïdes (T et B) et non-lymphoïdes (myéloïdes, érythroïdes et mégacaryoblastiques). Beaucoup de ces leucémies sont très complexes (leucémies mixtes ou biphénotypiques). Les leucémies mégacaryoblastiques sont particulièrement intéressantes car ces leucémies humaines sont associées à un mauvais pronostic et sont peu étudiées. L'étude a également révélée que la région *enhancer* du virus joue un rôle clé dans la pathogenèse, ceci correspondant à un plus grand nombre de leucémies lymphoïdes (T) induit par GV-1.2 en comparaison avec GV-1.4.

Une analyse phylogénétique détaillée a été effectuée grâce à l'obtention des séquences génomiques de GV-1.2 et GV-1.4. Cette étude a montré que le rétrovirus Graffi est très proche des virus SRS 19-6 et Moloney. L'étude phylogénique réalisée est très large, prenant en compte des membres de chaque classe de la famille des rétrovirus murins de la leucémies (écotropiques, amphotropiques, xénotropiques, polytropiques). Le rétrovirus Graffi, avec les virus SRS19-6, Moloney, Rauscher et Friend sont plus proches des virus de type Casistas que des virus xénotropiques. En outre, l'analyse a également révélée que l'enveloppe de HEMV, un prototype d'un virus ancestral et celle de CasBrE sont proches d'un point de vue phylogénique.

Finalement, afin d'utiliser les avantages du modèle 'Graffi', le profilage génique de chaque type de leucémies induites par ce rétrovirus a été effectuée grâce à la technologie des micropuces à haute densité (Affymetrix). De très nombreuses données sont issues de cette analyse et beaucoup de gènes non encore reliés à la leucémie ont été mis en évidence. Certains de ces gènes sont potentiellement des oncogènes et leurs fonctions méritent d'être analysées de façon plus approfondie. L'expression de certains gènes a été validée par RT-PCR.

En conclusion, le rétrovirus Graffi représente un modèle complexe mais très intéressant pour étudier divers types de leucémies.

L'expérience des micropuces a permis de faire ressortir des gènes dont il serait intéressant d'approfondir l'étude de la fonction. Globalement, l'utilisation des nombreuses données issues de l'analyse des micropuces permet une meilleure

compréhension des cellules leucémiques et des lignées hématopoïétiques correspondantes. Le profilage génique des leucémies induites par le rétrovirus Graffi a permis d'ouvrir sur de nombreuses perspectives, certaines à caractères thérapeutiques. Ainsi, à plus long terme, cela pourrait permettre de fournir des éléments s'ajoutant aux outils de diagnostic, et d'identifier des gènes ou voies de signalisation pouvant être ciblés par des molécules inhibitrices.

Mots-clés: rétrovirus murin de la leucémie, leucémie, oncogène, cancer, hématopoïèse

CHAPITRE I

ÉTAT DES CONNAISSANCES

1.1. Hématopoïèse et leucémie

1.1.1. Hématopoïèse

1.1.1.1. Définition

L'hématopoïèse (du grec, *haima*: sang et *poiesis*: fabrication) est un processus complexe et hautement régulé qui permet la production des globules rouges, des plaquettes et des globules blancs de l'organisme. Toutes ces cellules différenciées proviennent de mêmes cellules souches hématopoïétiques situées dans la moelle osseuse. Les cellules sanguines ont une durée de vie tellement courte qu'elles doivent être continuellement renouvelées à partir de cellules souches de la moelle osseuse par le processus de l'hématopoïèse. Chez l'humain, on estime la production quotidienne moyenne de 100 milliards de globules blancs, 200 milliards de globules rouges, et 100 milliards de plaquettes (Kaushansky, 2006).

1.1.1.2. Présentation générale des différentes cellules sanguines

Les globules rouges (érythrocytes) et les plaquettes (thrombocytes) sont de petites cellules énuclées, les unes impliquées dans le transport des gaz (oxygène et gaz carbonique) dans l'organisme, et les autres dans la coagulation du sang (plaquettes).

Les globules blancs (leucocytes) représentent plusieurs catégories distinctes de cellules. On distingue les globules blancs polynucléaires (basophiles, éosinophiles, neutrophiles), les monocytes (mononucléaires) et les lymphocytes (également mononucléaires). Ils sont impliqués dans la défense de l'organisme contre les différents pathogènes. Les

polynucléaires et les monocytes sont impliqués dans l'immunité innée et les lymphocytes dans l'immunité spécifique.

Les polynucléaires ou granulocytes ont un noyau segmenté et leur cytoplasme est rempli de granules (lysosomes) qui contiennent des enzymes capables de tuer ou de digérer les bactéries ou autres pathogènes. Les monocytes, comme les polynucléaires, se rendent jusqu'aux sites d'infection où ils se différencient en très grosses cellules, les macrophages. Les macrophages sont capables de phagocyter des cellules entières, comme des cellules infectées par des virus ou de gros pathogènes. De plus, les macrophages sont capables de présenter des peptides spécifiques de ces pathogènes aux autres cellules du système immunitaire (présentation de l'antigène avec le complexe majeur d'histocompatibilité I ou II). Un autre composant de l'immunité innée est le système du complément. Les macrophages sécrètent certains éléments du complément et ont aussi à la surface les récepteurs du complément qui permettent la phagocytose des pathogènes opsonisés.

Les lymphocytes B vont être à l'origine de la production d'une masse considérable d'anticorps lorsqu'ils reçoivent le signal de la présence d'un antigène spécifique. Les anticorps permettent l'activation du système du complément et permettent l'opsonisation des pathogènes pour faciliter leur phagocytose. Les lymphocytes T sont capables de détruire directement les cellules étrangères qu'ils rencontrent par cette reconnaissance spécifique d'antigènes. Les cellules T cytotoxiques tuent les cellules infectées par des pathogènes en injectant des granules lytiques et en induisant l'apoptose.

Deux autres sortes de cellules existent, moins nombreuses que les précédentes. Il s'agit des cellules dendritiques et des cellules NK (*Natural Killer*). Les cellules dendritiques sont une autre sorte de cellules présentatrices d'antigènes comme les macrophages et qui ont comme fonction première d'activer les cellules T par la présentation d'antigène. Les cellules NK sont capables de tuer les cellules infectées de la même façon que les cellules T cytotoxiques mais n'ont pas besoin d'être activées au préalable par le mécanisme de présentation d'antigène (Janeway et al. 2004).

Les cellules sanguines, le système immunitaire et l'hématopoïèse sont essentiels au fonctionnement des êtres multicellulaires qui peuvent être infectés par des pathogènes: ces systèmes sont conservés dans de nombreuses espèces et même chez les Invertébrés par exemple chez les Éponges. Chez la *Drosophile*, les cellules du système immunitaire

sont appelées hémocytes et sont composées de 3 types majeurs: plasmocytes hyalins, cellules cristallines et lamellocytes. Elles sont impliquées dans l'immunité innée et les plasmocytes hyalins sont notamment l'équivalent des monocytes/macrophages. Même les gènes importants dans l'hématopoïèse des Mammifères ont leur équivalent chez d'autres espèces. Ainsi, Serpent, le gène de la Drosophile homologue de Gata-1,-2,-3 est essentiel à la survie et différenciation des 3 groupes d'hémocytes (Hartenstein 2006). Étant donné la conservation du système entre les espèces, on peut comprendre que les observations obtenues chez la souris sont majoritairement vérifiables chez l'humain également.

1.1.1.3. Hématopoïèse embryonnaire

Le système hématopoïétique chez les Mammifères a comme origine embryonnaire le feuillet mésodermique. L'origine des cellules hématopoïétiques est commune avec celui du système vasculaire endothélial et la cellule progénitrice est nommée hémangioblaste. Ce système hématopoïétique doit se mettre en place très tôt lors du développement à cause de la nécessité du transport de l'oxygène par les globules rouges (revu dans Cumano et Godin , 2007). Ainsi, au stade E7.0-E7.5 (chez la souris), les premiers îlots sanguins, nommés îlots de Wolff et Pander, apparaissent au niveau du sac vitellin de l'embryon. Les cellules au centre de ces îlots sont des érythroblastes primitifs: ils contiennent de la globine embryonnaire et ne se différencient pas en cellules énuclées. Peu à peu, il y a formation de lacs vasculaires à la surface du sac vitellin, puis de vaisseaux sanguins. La circulation entre le sac vitellin et l'embryon s'effectue approximativement au stade E8. A partir de ce stade, il y a également présence dans le sac vitellin de progéniteurs myéloïdes (Cumano et al 1996, Palis et al. 1999).

Il existe un autre lieu de synthèse des progéniteurs hématopoïétiques qui est la région de la splanchnopleura para-ortique (Godin et al.1993, 1995) et qui devient au cours du développement la région AGM (aorte-gonade-mésonéphros). Elle contient des progéniteurs multipotents lymphoïdes/érythroïdes/myéloïdes (Cumano et al. 1996).

Avec le développement de l'embryon, l'activité hématopoïétique du sac vitellin décline très vite entre le stade E11 et E12 (chez la souris) pour faire place à une hématopoïèse hépatique. Hypothétiquement les cellules souches produites dans la région de la splanchnopleura para-ortique et dans le sac vitellin migreraient dans le foie fœtal dans lequel elles se différencieraient. Une érythropoïèse définitive est établie dans le foie au

stade E9.5. A E11.5, le foie devient le lieu de maturation des différentes lignées hématopoïétiques (érythroïde/lymphoïde/myéloïde).

Peu de temps avant la naissance, lorsque la croissance des os est assez avancée, l'hématopoïèse au niveau de la moelle osseuse se met en place. A la naissance, la moelle osseuse sera le site majoritaire de l'hématopoïèse, avec une hématopoïèse secondaire dans la rate (chez la souris uniquement et dans des cas pathologiques chez l'humain) et aucune dans le foie.

1.1.1.4. L'hématopoïèse

1.1.1.4.1. Processus général

Le processus général de l'hématopoïèse, qui est pyramidal, est illustré sur la figure 1.1 (Kaushansky 2006). Les cellules souches hématopoïétiques (HSC, *hematopoietic stem cells*) se trouvent au sommet de la pyramide. Ces cellules sont pluripotentes, dans un stade indifférencié, capables de se diviser indéfiniment et de donner naissance à tous les types de cellules sanguines. Pas à pas, certaines de ces cellules souches vont être entraînées dans des étapes de différenciation alors que certaines vont avoir comme rôle de maintenir le pool de HSC de la moelle osseuse. Les cellules dirigées dans le processus de différenciation deviennent des cellules progénitrices: ces cellules ont un nombre de division plus limité et sont programmées plus étroitement pour devenir une lignée cellulaire précise. Ainsi 2 types de cellules progénitrices se différencient à partir des HSCs: le progéniteur commun lymphoïde (CLP, *common lymphoid progenitor*) et le progéniteur commun myéloïde (CMP, *common myeloid progenitor*).

Le CLP donne naissance à un progéniteur commun aux cellules T et NK (NKTP) et le progéniteur commun aux cellules B (BCP). Les cellules B matures se différencient à partir de BCP. NKTP se différencie soit en progéniteur commun aux leucémies T (TCP) ou en précurseur des cellules NK (NKP). Les précurseurs continuent leur différenciation jusqu'à la maturation complète des cellules T, et NK. Il est à noter que les cellules T terminent leur différenciation dans le thymus.

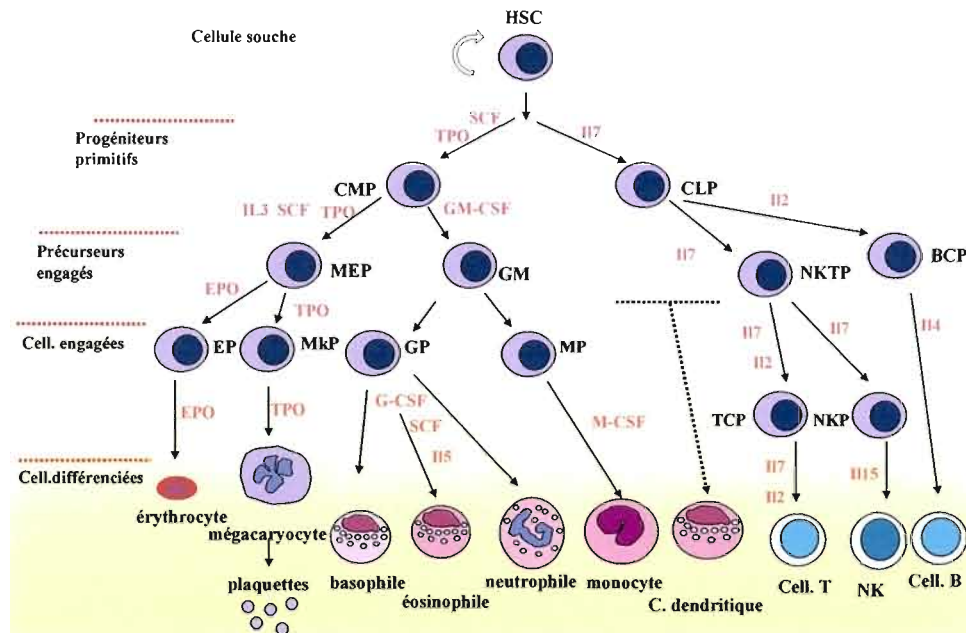


Figure 1.1: Schéma général de l'hématopoïèse (système pyramidal).

HSC: cellule souche hématopoïétique; CLP: progéniteur commun lymphoïde; TNK: progéniteur des cellules T et NK; BCP: progéniteur des cellules B; TCP: progéniteur des cellules T; NKP: précurseur des cellules NK; CMP: progéniteur commun myéloïde; MEP: progéniteur érythro-mégacaryocytaire; GM: progéniteur granulo-monocytaire; EP: précurseur érythroïde; MkP: précurseur mégacaryocytaire; GP: précurseur granulocytaire; MP: précurseur monocytique. Les facteurs de croissance les plus importants sont notés en rouge. EPO: érythropoïétine; TPO: thrombopoïétine. Figure adaptée de Kaushansky 2006 (Kaushansky 2006).

Le CMP donne naissance à la fois au progéniteur myéloïde (GM) qui va se différencier en granulocytes/monocytes/cellules dendritiques, et au progéniteur bipotent érythro-mégacaryocytaire (MEP) qui va se différencier en érythrocytes et en mégacaryocytes (qui produisent les plaquettes). GM peut se différencier en 2 voies distinctes pour séparer la voie des granulocytes polynucléaires (basophiles, éosinophiles, neutrophiles) et celle des monocytes (qui vont se différencier en macrophages).

Les cellules semblent être guidées tout au long du processus de l'hématopoïèse par des dialogues intercellulaires avec des messages émis entre les cellules sous forme de cytokines et facteurs de croissance. L'environnement cellulaire est donc déterminant pour la différenciation des cellules. Les cellules stromales de la moelle osseuse sont connues pour jouer un rôle très important dans ce processus.

1.1.1.4.2. Les cellules souches hématopoïétiques (HSCs) et la moelle osseuse

1.1.1.4.2.1. Les HSCs

Les HSCs sont les cellules les plus indifférenciées de toutes les cellules du système hématopoïétique et le pool de HSCs doit donc être soigneusement conservé tout au long de la vie de l'organisme. Les HSCs sont classées en 2 groupes: les LT-HSCs (*long-term* HSCs) et les ST-HSCs (*short-term* HSCs). Les LT-HSCs sont capables d'un nombre de division plus grand que les ST-HSCs; elles sont quiescentes alors que les ST-HSCs se divisent activement.

Afin que le pool de HSCs soit maintenu, les cellules se divisent soit asymétriquement soit symétriquement. La division asymétrique va donner 2 cellules filles, l'une qui reste une HSC et l'autre qui va suivre le chemin de la prolifération et de la différenciation. La division symétrique donne 2 cellules filles qui restent dans le pool des HSCs. Dans des conditions normales, les divisions asymétriques sont les plus fréquentes.

Les HSCs portent des marqueurs de surface spécifiques et qui les différencient des autres progéniteurs hématopoïétiques. Il est très important pour la purification et la manipulation de ces cellules de connaître le phénotype précis des HSCs. La fraction CD34⁺Lin⁻ des cellules de la moelle osseuse est très enrichie en HSCs. Mais des études plus poussées sont effectuées pour permettre une purification plus parfaite des LT-HSCs. Une étude récente montre que les HSCs seraient définies par le phénotype CD150⁺CD48⁻CD244⁻ alors que les phénotypes CD150⁻CD48⁺CD244⁻ ou CD150⁻CD48⁺CD244⁺ pourraient permettre la purification de progéniteurs plus différenciés (Kim et al. 2006).

Des HSCs se trouvent également dans la circulation sanguine et dans le sang de cordon ombilical et peuvent être transplantées pour reconstruire le système hématopoïétique.

1.1.1.4.2.2. Microenvironnement de la moelle osseuse

La moelle osseuse est formée dans la cavité de l'os. C'est une structure complexe, dont l'organisation est peu à peu découverte (figure 1.2). Il existe dans cette cavité des cellules stromales formant le microenvironnement de la moelle osseuse, et des vaisseaux sanguins. Dans la moelle osseuse se trouvent tous les progéniteurs hématopoïétiques qui sont retenus dans cet environnement jusqu'à ce que leur différenciation soit achevée, et ils passent alors dans la circulation sanguine. Autres que les cellules hématopoïétiques, il faut noter également la présence de cellules souches mésenchymateuses qui donnent naissance par exemple aux cellules stromales et aux cellules endothéliales.

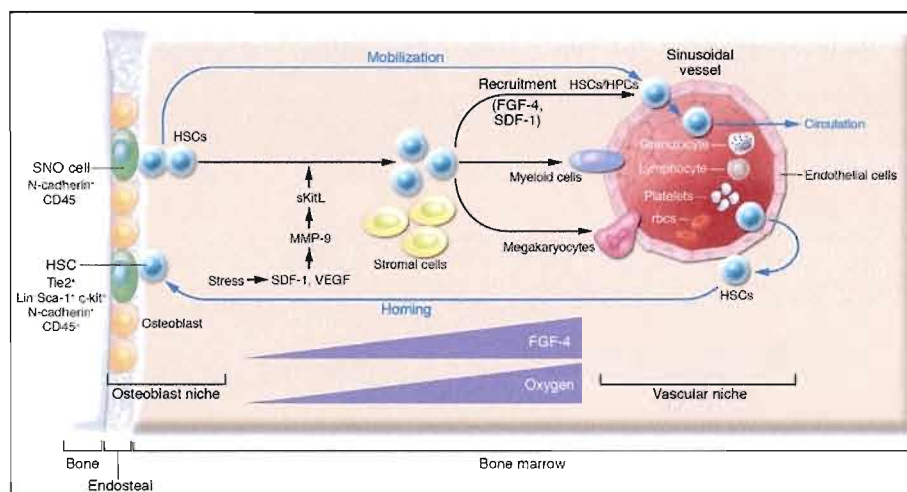


Figure 1.2: Organisation du microenvironnement de la moelle osseuse. Figure tirée de Yin et Li, 2006. La niche ostéoblastique et la niche vasculaire y sont représentées.

Deux structures dans la cavité osseuse sont particulièrement importantes et sont nommées la niche ostéoblastique et la niche vasculaire. Les niches sont les gardiens du pool de HSCs: les HSCs se regroupent dans ces sites et s'y auto-renouvellent. La fonction de ces niches est de réguler l'équilibre entre les divisions asymétriques et symétriques et les facteurs sécrétés par les cellules composant la niche assurent la survie des HSCs (Yi and Li 2006).

La niche ostéoblastique se situe le long de la paroi des os et la niche vasculaire est composée de cellules endothéliales qui composent la paroi des vaisseaux sanguins. De plus en plus de modèles penchent vers l'hypothèse selon laquelle les HSCs seraient plus

quiescentes dans la niche ostéoblastique alors qu'elles seraient plus prolifératives dans la niche vasculaire. La niche vasculaire fournit un environnement plus riche en nutriments, en facteurs de croissance, et en oxygène. Cette niche pourrait également faciliter la migration trans-endothéliale des cellules hématopoïétiques vers la circulation sanguine ainsi que faciliter le phénomène de *homing* (Yi and Li 2006).

Que ce soit entre les cellules hématopoïétiques et les cellules stromales, ou avec celles de la niche ostéoblastique ou vasculaire, ou même entre elles, les cellules s'envoient des messages chimiques sous formes de cytokines pour réguler leur prolifération ou différenciation. Des molécules d'adhésion spécifiques permettent aussi la rétention des cellules hématopoïétiques dans certains secteurs précis de la moelle osseuse.

1.1.1.4.2.3. Les facteurs de croissance

Certains facteurs de croissance sont nécessaires à la survie et à la prolifération des cellules multipotentes et cela à tous les stades de différenciation. Ces facteurs se trouvent dans le microenvironnement de la moelle osseuse. Ils se fixent sur des récepteurs spécifiques à la surface des cellules hématopoïétiques envoyant des signaux de survie, de prolifération ou de différenciation à l'intérieur de la cellule.

Ainsi Notch1 est exprimé à la surface des HSCs et Jag1 à la surface des ostéoblastes et des cellules stromales. Les signaux envoyés par le contact de ces 2 molécules permettent l'expansion de la population des HSCs (Yi and Li 2006). De même, l'interaction de ANG1 (angiopontine 1) située sur les ostéoblastes et TIE2 (une protéine tyrosine kinase) située sur les HSCs favorise la quiescence ainsi que l'adhérence sur la niche ostéoblastique des HSCs (Yi and Li 2006).

Certains des facteurs de croissance sont solubles (cytokines). Ils agissent sur les cellules hématopoïétiques à tous les stades de différenciation. Parmi ces facteurs, on peut citer les plus importants et les plus caractérisés: l'érythropoïétine (Epo), la thrombopoïétine (Tpo), le facteur *steel* (SCF pour Steel cell factor, il est le ligand du récepteur Kit), le ligand de Flt3, le G-CSF (*Granulocytic-Colony Stimulating Factor*), le GM-CSF (*Granulocytic-Monocytic CSF*), les interleukines, 2 (IL-2), 3 (IL-3) et 7 (IL-7) (Kaushansky 2006). SCF et la Tpo sont très importants pour les étapes précoces de la différenciation et notamment pour la survie et la prolifération des cellules souches

hématopoïétiques. La Tpo est aussi nécessaire à la mégacaryopoïèse et à la production plaquettaire. L'Epo agit sur l'érythropoïèse et son effet rapide sur l'augmentation de la quantité de globules rouges est bien connu par de nombreux sportifs professionnels. Le GM-CSF et le G-CSF sont importants pour la différenciation myéloïde. IL-7 agit plus spécifiquement sur la lignée lymphoïde (figure 1.1).

Les sites de production de ces facteurs solubles ne sont pas forcément le microenvironnement de la moelle osseuse. Ils arrivent dans la moelle osseuse par la circulation sanguine. Ainsi, l'Epo et la Tpo sont produits majoritairement par le foie et les reins, le G-CSF par les macrophages et les cellules endothéliales (Kaushansky 2006).

1.1.1.5 Décisions d'engagement dans une lignée et réseau de régulation génique

Un type cellulaire est unique par rapport à un autre grâce au programme génique qui est activé spécifiquement. C'est ce programme qui est modifié peu à peu lors des étapes de différenciation à partir des cellules souches hématopoïétiques.

Certains gènes, éteints dans les HSCs quiescentes vont être activés et vont provoquer la prolifération des cellules. Par exemple, l'expression de Kit dont le ligand est SCF (voir section précédente) maintient les cellules dans un état prolifératif et c'est le complexe de SCL avec ses partenaires d'interactions qui agit en activant la transcription de Kit (Loose et al. 2007).

L'activation ou l'inhibition de certains gènes clés vont guider les cellules dans une voie de différenciation précise. Par exemple, une augmentation de l'expression de Gata1 dans les cellules souches va les conduire vers la différenciation myéloïde (Loose et al. 2007).

Le programme génique semble être régulé par les signaux extracellulaires et par l'action des facteurs de transcription à l'intérieur de la cellule.

Ainsi une hypothèse pour expliquer le destin d'une cellule est qu'il résulterait d'une compétition entre plusieurs facteurs de transcription dans une cellule en voie de différenciation (Loose et al. 2007). Par exemple, 2 facteurs de transcription nommés *a* et *b* auraient des rôles antagonistes: ils vont se fixer sur les mêmes sites mais n'auront pas le même effet sur l'expression des gènes cibles. Si dans une cellule, le facteur *a* se fixe

avant b , il va déclencher une série de réaction amenant la cellule à se différencier dans une voie; si le facteur b se fixe avant a , une autre voie de différenciation sera engagée.

Par exemple, une forte expression de Scl va plutôt conduire les cellules souches à se différencier en cellules myéloïdes (CMP) plutôt qu'en cellules lymphoïdes (CLP). Cependant, la protéine SCL agit avec des partenaires d'interaction. Un de ses partenaires d'interaction est E2A et si la concentration de celui-ci augmente, alors la différenciation se fera plutôt vers la lignée lymphoïde (Loose et al. 2007).

De même, il existe une compétition entre les facteurs de transcription GATA1 et PU.1 dans les cellules CMP (figure 1.3). Les 2 facteurs sont antagonistes, c'est-à-dire que GATA1 inhibe l'activité de PU.1 et PU.1 inhibe l'activité de GATA1, et cela par interaction physique entre les 2 protéines. GATA1 dirige les CMP vers la différenciation en MEP (progéniteur érythro-mégakaryocytaire) alors que PU.1 dirige vers la différenciation en GM (progéniteur granulocytaire/monocytaire). Ainsi, une concentration plus grande en GATA1 ou en PU.1 va déterminer le destin de la cellule (Graf 2002).

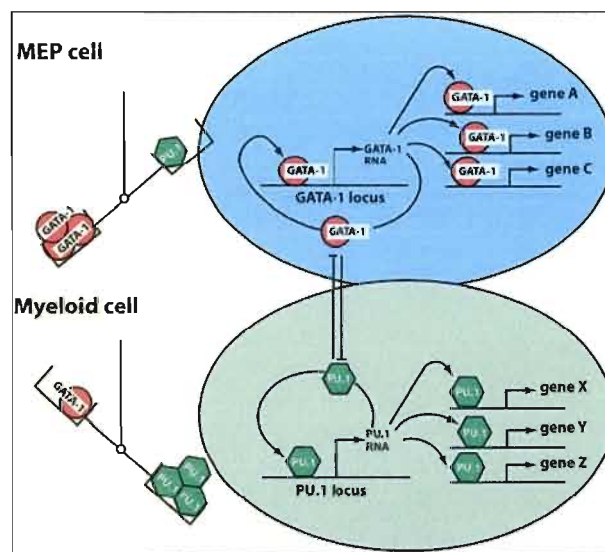


Figure 1.3: Antagonisme entre GATA-1 et PU.1. Figure tirée de Graf, 2002. MEP: progéniteur érythro-mégakaryocytaire.

1.1.1.6. Les différentes lignées

Une description plus détaillée sur la différenciation des principales lignées va être présentée avec une attention particulière portée sur les facteurs de transcription ou gènes clés de ces lignées.

1.1.1.6.1 Différenciation de la lignée lymphoïde

La lignée lymphoïde débute à partir du CLP (*common lymphoid progenitor*). Les caractéristiques de l'engagement d'une cellule dans la voie lymphoïde sont l'augmentation des gènes Flt-3, IL7R α (récepteur de l'interleukine 7) et des gènes de recombinaison Rag1 et Rag2. Deux facteurs de transcription, Ikaros et PU.1, ont une implication majeure dans l'initiation de la différenciation lymphoïde. Ainsi, les souris *knock-out* pour Ikaros ne produisent aucune cellule T, B ou NK. Les progéniteurs lymphoïdes les plus primitifs existent tout de même dans ces souris *knock-out* mais n'expriment pas les gènes Flt-3, IL7R α et Rag1,-2. Les souris *knock-out* pour Spi1 (PU.1) meurent environ au moment de la naissance et ne contiennent aucune cellule T, B, et NK et ont un nombre très réduit de progéniteurs lymphoïdes (Ye and Graf, 2007).

1.1.1.6.1.1. Développement des cellules B

Les étapes clés du développement des cellules B sont décrites par rapport à l'évolution des réarrangements des chaînes de l'immunoglobuline qui vont servir à établir les récepteurs des cellules B (BCRs: *B cell receptors*). Les différentes étapes du développement des cellules B sont illustrées sur la figure 1.4. Le récepteur des cellules B sert à reconnaître les antigènes. Il est composé d'une chaîne lourde et d'une chaîne légère d'immunoglobuline. Chaque chaîne d'immunoglobuline provient du réarrangement de 3 segments, V, D et J. Comme de nombreuses copies existent pour chaque segment, le réarrangement des chaînes est légèrement différent pour chaque cellule B formée, ce qui garantit la diversité des BCRs et permet à l'organisme de répondre à de nombreux pathogènes (Janeway et al. 2004).

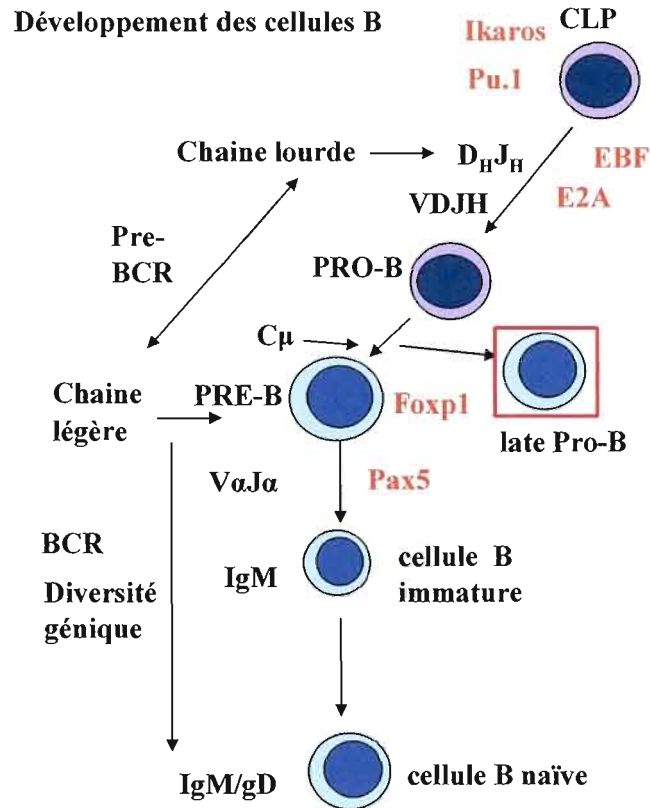


Figure 1.4: Le développement des cellules B. Illustration des étapes principales à partir du progéniteur commun lymphoïde (CLP). Les facteurs de transcription clés sont indiqués en rouge. Le stade qui correspond aux leucémies induites par le rétrovirus Graffi est également indiqué en encadré rouge (chapitre IV).

Lors du développement des cellules B, le premier stade, appelé pro-B, est caractérisé par le réarrangement des segments D_H et J_H suivi de celui des segments V_H et DJ_H de la chaîne lourde de l'immunoglobuline. La fin du réarrangement VDJ_H conduit à l'expression de la chaîne μ de l'immunoglobuline et ceci atteste du passage du stade pro-B au stade pré-B avec l'expression du pré-BCR à la surface des cellules. Il s'ensuit le réarrangement de la chaîne légère de l'immunoglobuline pour finalement conduire à l'expression d'un BCR fonctionnel à la surface. Les cellules B immatures, exprimant des molécules d'IgM à la surface migrent dans la rate et si elles sont tolérantes au soi, elles continuent leur différenciation en cellules B naïves (Janeway et al. 2004).

Des facteurs de transcription contrôlent les étapes clés du développement des cellules B (figure 1.4) et des points de contrôle sont situés après les différentes étapes des réarrangements. Le rôle des facteurs de transcription principalement impliqués dans le développement des cellules B va être discuté brièvement.

Le facteur de transcription PU.1 qui a un rôle dans l'engagement des cellules dans la voie lymphoïde, aiderait également l'engagement dans la voie spécifique du développement des cellules B: il aiderait à l'augmentation de l'expression des gènes *Ebf* (*early B-cell factor*) et *IL7ra*. En effet l'expression forcée de *Spi1* (PU-1) dans des cellules *knock-out* pour ce gène a comme conséquence une augmentation de l'expression de *Ebf* et *IL7ra*. De plus, ces 2 gènes ont des sites de liaison à PU.1 dans leurs régions régulatrices (Ye et Graf, 2007).

Les facteurs E2A et EBF ont un rôle dans l'initiation des réarrangements des segments D_H et J_H . En l'absence de ces facteurs, le développement est arrêté au stade pro-B avant les étapes de réarrangement. L'expression forcée de *Ebf* peut activer le programme de différenciation en cellule B même en l'absence de PU.1 ou E2A. Le rôle de EBF est donc en aval de celui de PU.1 et E2A: des sites de liaisons de E2A sur les séquences régulatrices de EBF ont effectivement été trouvés. L'expression de *Ebf* est également activée par la voie de signalisation passant par l'interleukine 7 (*IL7ra* et *Stat5*). EBF stimule l'expression de *Mb1* (CD79a, qui permet l'escorte du BCR à la surface) et de *Pax5* (Ye et Graf, 2007).

Pax5 agit principalement après le stade pro-B. Il a un double rôle, celui d'activer les gènes spécifiques de la lignée B et celui de réprimer des gènes de totipotence ou d'autres lignées. Les cellules B *knock-out* pour *Pax5* sont arrêtées au stade pro-B et ont une grande capacité d'auto-renouvellement et de totipotence comme les progéniteurs hématopoïétiques. Les gènes spécifiques ciblés par *Pax5* sont ceux constituant le pré-BCR et le BCR. Les gènes réprimés par *PAX5* sont par exemple *Notch-1* (qui conduit à la différenciation des cellules T) et *c-fms* (qui conduit à la différenciation de la lignée myéloïde). Il réprime également *Flt3* (Fuxa and Skok, 2007).

Finalement, *Foxp1* est également un facteur clé dans le développement des cellules B. Il agit dans la transition du stade pro-B vers pré-B, lors de la recombinaison V(D)J (Fuxa and Skok, 2007).

1.1.1.6.1.2. Développement des cellules T

La première étape importante pour le développement des cellules T est la migration d'un *progéniteur multipotent* de la moelle osseuse vers le thymus. Des études récentes indiquent que le phénotype de ces cellules multipotentes porte le phénotype Lin⁻Sca-1⁺Kit⁺ (Ye et Graf, 2007). Le thymus n'offre pas le même microenvironnement que la moelle osseuse et les signaux extracellulaires comme ceux envoyés par les ligands de Notch1 ou par l'interleukine 7 vont amener ces cellules multipotentes vers la différenciation en cellules T.

Les principales étapes de différenciation des cellules T sont illustrées sur la figure 1.5. Les cellules pro-T sont d'abord caractérisées par un phénotype double-négatif (DN) qui correspond à l'absence d'expression des molécules CD4 et CD8 (Janeway et al. 2004).

Les étapes de différenciation sont marquées, comme pour le développement des cellules B, par le réarrangement progressif du récepteur des cellules T (TCR). Il existe 2 lignées distinctes de cellules T portant soit un TCR composé des chaînes δ et γ soit composé des chaînes α et β . Le récepteur des cellules CD4 et CD8 est composé des chaînes α et β associées aux différentes chaînes de la molécule CD3. Le réarrangement allélique des chaînes δ , γ , β se passe au niveau du stade DN. Le premier point de contrôle est le contrôle du réarrangement réussi des chaînes δ , γ , β (Janeway et al. 2004).

Le point de contrôle vérifiant le réarrangement de la chaîne β est hautement régulé et est nommé la sélection β . Les cellules n'ayant pas réussi le réarrangement meurent par apoptose. A partir de ce point, les cellules ayant réussi passent du stade DN à DP (double positif) avec l'expression concomitante de CD4 et de CD8. A partir de ce moment, les cellules commencent le réarrangement de la chaîne α du TCR. Une fois ce réarrangement réalisé, un autre point de contrôle, appelé sélection positive, teste les cellules pour la tolérance au soi. De nombreuses cellules échouent ce test et meurent par apoptose. Les cellules survivantes continuent leur différenciation en cellules simple-positives, soit en CD4⁺CD8⁻ ou en CD4⁻CD8⁺ (Janeway et al. 2004).

Les facteurs de transcription majeurs impliqués dans la différenciation des cellules T sont NOTCH1, GATA3, E2A et HEB. Il faut noter qu'une diminution de l'expression de PU.1 est nécessaire pour l'engagement dans la voie de différenciation. Certains gènes spécifiques des cellules T comme Rag-1, Rag-2, les chaînes γ et ϵ du CD3, Lat et Lck doivent augmenter leur expression dès les premières étapes de la différenciation.

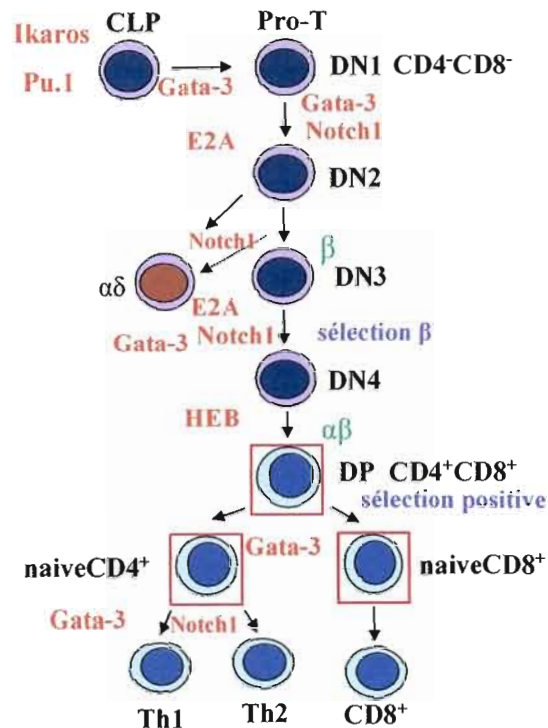


Figure 1.5: Le développement des cellules T. Illustration des étapes principales à partir du progéniteur commun lymphoïde (CLP). Les facteurs de transcription clés sont indiqués en rouge. Le stade qui correspond aux leucémies induites par le rétrovirus Graft est également indiqué en encadré rouge (chapitre IV).

1.1.1.6.1.2.1. Notch1

Le facteur de transcription Notch1 est indispensable au développement des cellules T. Lors d'un *knock-out* de Notch1, les précurseurs ne peuvent pas se développer en cellules T dans le thymus et, de façon étonnante, des cellules B se développent à la place. L'expression forcée de Notch1 dans des précurseurs hématopoïétiques conduit à la différenciation en cellules T et bloque en même temps le programme de différenciation en cellules B. NOTCH1 et les gènes qu'il active directement sont les plus fortement exprimés dans les stades DN avant la sélection β . Après ce stade, l'expression de Notch1 diminue.

Cependant, Notch1 joue également un rôle important dans les étapes de différenciation plus en aval, soit au moment de la sélection β , de la sélection positive et de la différenciation des cellules $CD4^+$. Ainsi son rôle dans le réarrangement du TCR β a été montré par un *knock-out* conditionnel avec l'activation de la délétion dans les cellules DN intra-thymiques (stade correspondant au réarrangement de la chaîne β). L'expression forcée et constitutive de ce gène perturbe l'expression des gènes spécifiques au stade DP, perturbant la sélection positive, et pouvant conduire au développement de la leucémie (Rothenberg and Taghon, 2005)

1.1.1.6.1.2.2. Gata3

Gata3 est le facteur de transcription unique qui apparaît spécifiquement dans une seule lignée hématopoïétique, soit la lignée T (Rothenberg and Taghon, 2005). Il faut noter qu'il intervient tout de même dans des cellules non-hématopoïétiques. Les souris *knock-out* pour Gata3 meurent en cours de gestation; il y a une réduction du nombre de progéniteurs hématopoïétiques avec une absence complète de maturation de la lignée T et un défaut de maturation des cellules NK. Les gènes cibles de GATA3 sont notamment Rag-2, IL4 et les gènes favorisant la formation du TCR.

L'expression de Gata3 augmente quand la différenciation en cellules T commence et atteint un pic lors de la prolifération intense qui se produit au stade de la sélection β . Des expériences de *knock-outs* conditionnels prouvent ce rôle dans la transition du stade DN au stade DP (Rothenberg and Taghon, 2005).

Plus en aval dans le processus de différenciation, l'expression de Gata3 est également forte lors de la sélection positive des cellules $CD4^+$ et il bloque le développement des cellules $CD8^+$ (Rothenberg and Taghon, 2005).

De façon inattendue et différente des autres facteurs de transcription clés, l'expression forcée de Gata3 dans des précurseurs hématopoïétiques non-thymiques ou thymiques ne permet pas d'influencer la différenciation vers la lignée T. Même au contraire, cela abolit l'engagement vers la lignée T (Rothenberg and Taghon, 2005). Il est possible que GATA3 ait un rôle diversifié et que ses actions sur la différenciation soient réalisées par exemple en synergie avec des co-facteurs bien spécifiques.

1.1.1.6.1.2.3. E2A, Heb, E2.2

Des facteurs de transcription de la famille hélice-boucle-hélice de type I, les facteurs E, jouent également un rôle fondamental dans la différenciation des cellules T. La perte de ces facteurs abolit complètement le développement des cellules T. Les thymocytes n'expriment pas seulement le facteur E2A également impliqué dans le développement des cellules B mais aussi HEB et E2.2 (Tcf4). L'hétérodimère E2A/HEB agit comme un régulateur positif de Gata3. Certains des gènes ciblés par ces facteurs de transcription E sont notamment Rag-1, Rag-2 et Cd4. E2A agit dans les stades plus précoces par rapport à HEB et E2.2. HEB agit notamment dans la transition entre le stade DN et le stade DP (figure 1.5). Ces facteurs agissent en provoquant un arrêt du cycle cellulaire lors des différents points de contrôle rencontrés lors de la différenciation (Rothenberg and Taghon, 2005)

En conclusion, il est à noter que la cascade de régulation, c'est-à-dire l'ordre dans lequel agissent les différents facteurs de transcription impliqués dans la différenciation T, n'est pas aussi clairement établie que pour les cellules B, peut-être parce que ces facteurs de transcription jouent tous des rôles à de multiples étapes de la différenciation.

1.1.1.6.2. Différenciation de la lignée non-lymphoïde

La lignée non-lymphoïde se sépare en 2 branches à partir du CMP, l'une conduisant à la lignée myéloïde à partir du progéniteur granulo-monocytaire (GM) et l'autre aux lignées érythroïde et mégacaryocytaire à partir du progéniteur MEP (figure 1.1). Les CMPs se trouvent dans la fraction $\text{II7R}\alpha^{-}\text{Kit}^{+}\text{Lin}^{-}\text{Sca-1}^{-}$ des cellules de la moelle osseuse. Dans cette population, les cellules $\text{Fc}\gamma\text{II/III}^{\text{hi}}\text{CD34}^{+}$ définissent les GMs alors que les cellules $\text{Fc}\gamma\text{II/III}^{\text{lo}}\text{CD34}^{-}$ définissent les MEPs.

1.1.1.6.2.1. La lignée myéloïde

Le facteur de transcription PU.1 oriente les CMPs afin qu'ils prennent la voie de différenciation myéloïde. Les progéniteurs GM purifiés et mis en culture génèrent majoritairement des neutrophiles, des monocytes et macrophages et une population mineure d'éosinophiles, basophiles et cellules mastocytaires. Dans la moelle osseuse, un progéniteur éosinophilique a été isolé et il est caractérisé par l'expression d'interleukine 5. Il existe un progéniteur bipotent basophile/mastocytaire ainsi qu'un progéniteur bipotent monocyte/cellule dendritique.

Les facteurs de transcription impliqués majoritairement dans la lignée myéloïde sont PU.1, C/EBP α , IRF8, GFI1. Ces facteurs de transcription régulent l'expression de beaucoup de gènes myéloïdes comme ceux codant pour M-CSF, G-CSF, GM-CSF ou les composants des granules (lactoferrine) (Rosenbauer and Tenen 2007).

1.1.1.6.2.1.1. PU.1

PU.1 est un facteur de transcription de type Ets et est codé par le gène Spi1. Il est nécessaire à l'établissement de l'engagement dans la voie myéloïde et de l'expression des premiers gènes spécifiques de cette lignée. Des motifs de liaison à PU.1 sont présents sur les régions régulatrices de pratiquement tous les gènes myéloïdes. Spi-1 est exprimé fortement par les progéniteurs myéloïdes mais aussi par les cellules matures myéloïdes. Le *knock-out* de Spi-1 conduit à une hématopoïèse fœtale défectueuse avec une absence complète de cellules B et de macrophages (Rosenbauer and Tenen 2007).

PU.1 agit au niveau des stades précoces de la myélopoïèse mais aussi dans les stades tardifs comme la granulopoïèse. Les effets de PU.1 sont dose-dépendants: de fortes quantités de PU.1 supportent la production de macrophages alors que de faibles quantités conduisent plutôt à la production de granulocytes (Rosenbauer and Tenen 2007).

1.1.1.6.2.1.2. C/ebp α

Alors que PU.1 est nécessaire à la formation des CMP à partir des cellules souches, C/ebp α est nécessaire à la formation des GMs à partir des CMPs. C/ebp α est exprimé par les CMPs, GPs, les granulocytes mais pas par les macrophages. Le *knock-out* de ce gène conduit à une absence de différenciation myéloïde à partir des progéniteurs CMPs.

De façon surprenante, C/ebp α n'est pas nécessaire au développement de la granulopoïèse après le stade GM puisqu'elle s'effectue normalement dans des souris portant un *knock-out* conditionnel. C/EBP α régule la transcription de nombreux gènes mais joue également un rôle dans l'arrêt du cycle cellulaire à divers points de contrôle (Rosenbauer and Tenen 2007).

1.1.1.6.2.1.3. Irf8

La décision d'une cellule de se différencier en monocytes/macrophages ou en granulocytes est orchestrée par PU.1 à l'aide notamment d'un de ses partenaires d'interaction IRF8. Dans la lignée myéloïde, Irf8 est exprimé par les progéniteurs myéloïdes et les macrophages mais pas par les granulocytes.

Le *knock-out* de Irf8 conduit à l'apparition d'un syndrome myéloprolifératif chez la souris qui ressemble aux leucémies chroniques myéloïdes de l'humain. Dans les souris *knock-out*, le nombre de granulocytes et les précurseurs correspondants sont augmentés mais il y a moins de macrophages. Lorsque l'on restaure l'expression de ce gène, la différenciation en macrophages apparaît de nouveau (Rosenbauer and Tenen 2007).

1.1.1.6.2.1.4. Gfi1 et C/ebp ϵ

Gfi1 et C/ebp ϵ sont 2 facteurs de transcription qui agissent en aval des progéniteurs GMs et orientent la différenciation vers la granulopoïèse.

Gfi1 est un répresseur de la transcription. Son expression est élevée dans les neutrophiles. Son *knock-out* induit une hématopoïèse défectueuse avec une absence totale de neutrophiles. La différenciation des neutrophiles est bloquée juste en aval du stade GM (Rosenbauer and Tenen 2007).

Le *knock-out* de *C/ebpε* induit un phénotype semblable mais plus large, avec un blocage de toute la granulopoïèse et des macrophages matures anormaux (Rosenbauer and Tenen 2007).

1.1.1.6.2.2. La lignée érytho-mégacaryopoïétique

1.1.1.6.2.2.1. Différenciation érythropoïétique et mégacaryopoïétique

Ces 2 lignées sont très proches. Elles partagent beaucoup de facteurs de transcription en commun. Cependant elles finissent par devenir des cellules matures très spécialisées et très différentes et n'ont pas beaucoup de protéines en commun. Le progéniteur bipotent MEP a été isolé seulement récemment (Debili et al. 1996, Akashi et al. 2000, Vanucchi et al 2000). Auparavant, on pensait que les 2 voies déviaient à partir du progéniteur commun myéloïde (CMP). Ainsi, la découverte du MEP apporte une preuve supplémentaire de la promiscuité entre ces 2 lignées.

La mégacaryopoïèse conduit à la production de cellules géantes, les mégacaryocytes. La figure 1.6 schématise les étapes de la mégacaryopoïèse. Cette population mégacaryocytaire est peu nombreuse, représentant moins de 1% des cellules myéloïdes de la moelle osseuse (Patel et al. 2005). Cependant un mégacaryocyte est capable de produire à lui seul des milliers de plaquettes par un mécanisme nommé thrombopoïèse. La différenciation des mégacaryoblastes a pour but de produire une cellule de plus en plus grosse, dont le noyau devient polyploïde par le processus de l'endomitose. Le noyau polyploïde sert à synthétiser rapidement toutes les protéines qui vont être nécessaires à la formation et au fonctionnement des plaquettes. En même temps, il se forme un système complexe de membranes internes. Ce système va permettre la formation des membranes des nombreuses plaquettes à partir d'un seul mégacaryocyte. Lors de la thrombopoïèse, les plaquettes sont assemblées à partir des pseudopodes formés dans les mégacaryocytes matures (revu par Shivdasani et al 2001). La figure 1.6 montre une photographie de la formation des pseudopodes et des proplaquettes (Ishida et al. 1996). Les plaquettes sont des cellules anucléées qui contiennent des granules denses et quelques ARNs messagers (ceux des gènes *GpIba*, *GpIIb*, et *Pf4*) ainsi que la machinerie nécessaire à la traduction. Le mégacaryocyte et les plaquettes restent à l'heure actuelle des cellules mystérieuses, dont le rôle ne doit pas être simplifié à la coagulation du sang (voir chapitre IV)

Lors de l'érythropoïèse, les cellules en cours de différenciation deviennent de plus en plus petites. Les différentes étapes de l'érythropoïèse sont illustrées sur la figure 1.7. Peu avant la différenciation terminale, les globules rouges commencent à perdre leur noyau (réticulocytes). Ainsi les globules rouges sont chez les Mammifères des cellules énuclées, d'où leur durée de vie courte. Comme les plaquettes, quelques espèces d'ARNs messagers, principalement ceux de la globine, restent dans les globules rouges.

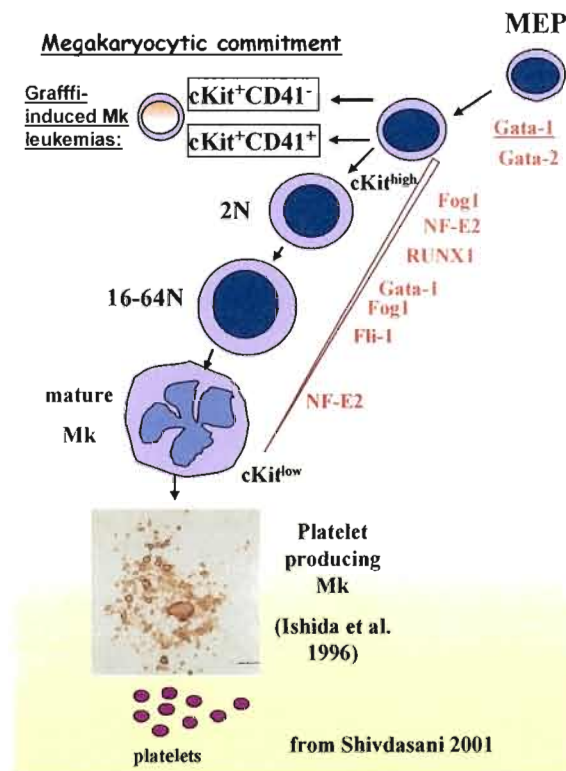


Figure 1.6: La mégacaryopoïèse. Illustration des étapes principales à partir du progéniteur érythro-mégacaryopoïétique (MEP). Les facteurs de transcription clés sont indiqués en rouge. Le stade qui correspond aux leucémies induites par le rétrovirus Graffi est également indiqué (chapitre IV).

Certains marqueurs de surface différencient les mégacaryoblastes et les érythroblastes (figures 1.6 et 1.7). Les 2 lignées expriment fortement Kit à des stades immatures mais l'expression de celui-ci diminue avec la différenciation. CD41 (glycoprotéine IIb) caractérise la lignée mégacaryoblastique: plaquettes, mégacaryocytes et

mégacaryoblastes (Uzan et al. 1991). TER119 et CD71 caractérisent la lignée érythroïde (Kina, et al. 2000, Lok and Loh. 1998).

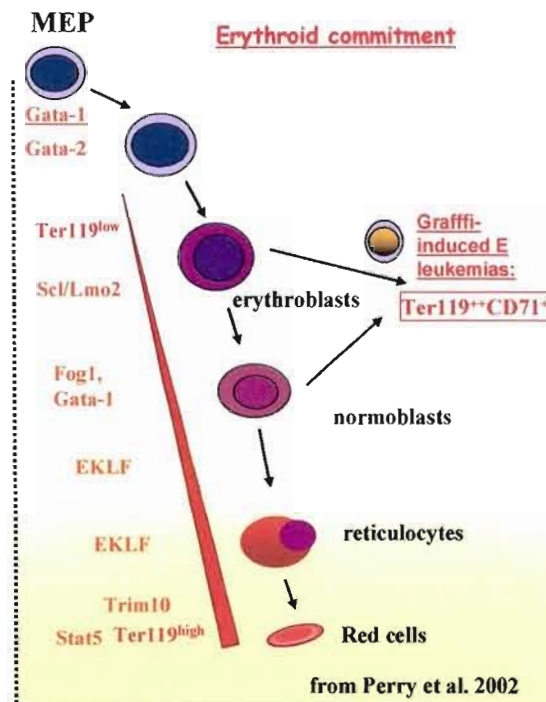


Figure 1.7: L'érythropoïèse. Illustration des étapes principales à partir du progéniteur érythro-mégacaryopoïétique (MEP). Les facteurs de transcription clés sont indiqués en rouge. Le stade qui correspond aux leucémies induites par le rétrovirus Graffi est également indiqué (chapitre IV).

1.1.1.6.2.2.2. Rôle clé de certains facteurs de transcription

Certains facteurs ayant des rôles clés dans la différenciation érythroïde ou mégacaryocytaire sont montrés sur les figures 1.6 et 1.7. Beaucoup de facteurs de transcription ont des rôles déterminants à la fois dans l'érythropoïèse et dans la mégacaryopoïèse (GATA1, FOG-1, SCL, GATA2, FLI-1). Certains autres ont des rôles plus spécifiques à une lignée et les exemples de EKLF et HERF1 (érythroïdes), RUNX1 et NFE2 (mégacaryocytaires).

1.1.1.6.2.2.2.1 Gata1

Une protéine jouant indubitablement un rôle dans l'érythropoïèse est GATA1. Elle appartient à la famille des facteurs de transcription en doigt de zinc. Gata1 a été premièrement identifié par sa capacité à lier des séquences régulatrices des gènes de la globine (Cantor et al. 2002). Depuis, des motifs de liaison à GATA1 ((T/A)GATA(A/G)) ont été identifiés au niveau de promoteurs et/ou *enhancers* de tous les gènes érythroïdes étudiés (Cantor et al. 2002). Des souris *knock-out* mâles hémizygotes pour Gata1 meurent au jour E10.5 de gestation d'une anémie sévère avec un arrêt de la maturation érythroïde au stade pro-érythroblastes (Gata1 est situé sur le chromosome X) (Orkin 2002). Les cellules souches embryonnaires Gata1⁻ ne sont pas capables de se différencier en culture au-delà du stade érythroblaste et sont conduites en apoptose, indiquant un rôle de Gata1 à la fois sur la survie cellulaire et aussi sur la maturation (Weiss and Orkin 1995).

Gata1, présent dans les mégacaryocytes, joue également un rôle fondamental dans la mégacaryopoïèse (Shivdasani et al. 2001). Comme dans l'érythropoïèse, les régions régulatrices de certains gènes spécifiques des mégacaryocytes comportent toutes des sites GATA quelquefois combinés à des sites Ets comme c'est le cas par exemple pour les gènes PF4 (*platelet factor 4*), p45/NFE2 et GpIIb (Shivdasani et al. 2001) et Fli-1 (Barbeau et al 1999). Les cellules mégacaryoblastiques issues de souris *knock-out* pour Gata1 montrent une maturation incomplète (Orkin 2002). Un *knock-out* sélectif de Gata1 dans les cellules mégacaryopoïétiques, conduit à la production de mégacaryocytes anormaux, petits et immatures et hyperprolifératifs (Orkin 2002, Shivadasani et al. 2001). La prolifération anormale de ces mégacaryocytes suggère que Gata1 est un régulateur négatif de la prolifération cellulaire dans les progéniteurs mégacaryopoïétiques et différent de celui joué dans la lignée érythropoïétique. GATA-1 favorise la maturation des mégacaryocytes et la formation des plaquettes (Orkin 2002, Shivadasani et al. 2001).

1.1.1.6.2.2.2.2. Fog-1

Fog-1 est également un facteur de transcription en doigt de zinc. FOG-1 a premièrement été identifié comme un facteur d'interaction avec GATA1 (Tsang et al. 1997). Fog-1 est exprimé de façon abondante dans les cellules érythroïdes et les cellules

mégacaryopoïétiques. Le *knock-out* de Fog-1 est similaire à celui de Gata1 (Tsang et al. 1998), c'est-à-dire létal en milieu de gestation du à une anémie sévère avec absence de maturation des érythroblastes. Cependant, contrairement aux souris Gata1^{-/-}, les souris Fog-1^{-/-} ne produisent pas du tout de cellules mégacaryopoïétiques (Cantor et al. 2002), ce qui suggère que Fog-1 a peut-être aussi un rôle indépendant de Gata1 dans la mégacaryopoïèse et que ce dernier interviendrait dans des étapes précoces de ce processus. FOG-1 peut jouer un rôle de répresseur ou d'activateur des actions de GATA1. Par exemple, il agit en synergie avec GATA1 pour activer le promoteur de p45/Nfe2 (Tsang et al. 1997), il réprime les activités de GATA1 sur le promoteur de Eklf et sur le récepteur 2 de la transferrine (Kawabata et al. 2001).

1.1.1.6.2.2.2.3. Scl et Lmo2

Scl (Tal1) est un facteur de transcription de la famille bHLH (*basic helix loop helix*). Il lie l'ADN sur des sites CANNTG, séquences consensus retrouvées sur des éléments régulateurs de gènes érythroïdes. Il est détecté dans les progéniteurs hématopoïétiques, dans la lignée érythroïde, dans les mégacaryocytes, et mastocytes. Il interagirait avec plusieurs co-facteurs pour agir sur les éléments régulateurs et promoteurs des gènes (Cantor et al 2002). Par exemple, SCL régule le gène glycophorine A (Ter119) spécifique de la lignée érythroïde (Lahlil et al. 2004).

L'implication de Scl dans la mégacaryopoïèse a été démontrée par un *knock-out* conditionnel de Scl. La délétion de Scl dans les souris adultes perturbe l'érythropoïèse ainsi que la mégacaryopoïèse: une absence des progéniteurs érythroïdes et mégacaryocytaires est observée avec cependant une production normale de cellules myéloïdes (monocytes et polynucléaires) (Hall et al. 2003).

Un des co-facteurs de SCL est notamment LMO2. Des souris *knock-out* pour Lmo2 meurent en cours de gestation d'anémie sévère avec une absence d'hématopoïèse primitive (Perry et al. 2002). Lmo2 aurait avec Scl des rôles plus spécifiques dans certaines lignées hématopoïétiques notamment érythroïdes. SCL et LMO2 forment un complexe avec E2A, LDB1 et GATA1 qui activerait des gènes érythroïdes (Perry et al. 2002).

1.1.1.6.2.2.2.4. Gata2

Son expression forcée dans les précurseurs érythroïdes induit la prolifération et bloque la différenciation. Les souris *knock-out* pour Gata2 meurent durant la gestation, dû à une sévère anémie apparaissant dès les premiers stades de l'hématopoïèse fœtale (Perry et al. 2002). L'expression de Gata2 précède celle de Gata1 lors des étapes de l'hématopoïèse et il faut une réduction de l'expression de Gata2 pour que la différenciation érythroïde puisse s'effectuer (Perry et al. 2002).

Gata2 est exprimé fortement par les mégacaryocytes. Des essais de différenciation hématopoïétiques de cellules embryonnaires combinés à une surexpression de Gata2 montrent que Gata2 dirige les cellules vers une différenciation érythroïde et mégacaryocytaire. La différenciation érythroïde se produit si Gata2 est exprimé au jour 6 de l'essai de différenciation et la différenciation mégacaryocytaire se produit au jour 5, indiquant que le phénomène est dépendant de l'état de différenciation des cellules (Kitajima et al. 2006).

Il est difficile de discerner la différence des rôles joués par Gata1 et Gata2 et certains émettent l'hypothèse qu'ils pourraient jouer des rôles redondants notamment dans la lignée mégacaryocytaire dans laquelle Gata1 et 2 sont tous deux fortement exprimés (Shivdasani et al, 2001).

1.1.1.6.2.2.2.5. Fli-1

Fli-1 est un facteur de transcription de la famille Ets. Il a premièrement été identifié dans des leucémies érythrocytaires causées par le virus de Friend. Une surexpression de Fli-1 inhibe la différenciation érythroïde et une répression de Fli-1 par l'érythropoïétine permet la réinduction du processus de différenciation. (Perry et al. 2002).

Les souris *knock-out* pour Fli-1 meurent en cours de gestation. Elles souffrent d'hémorragie et de défauts vasculaires. De plus, elles souffrent d'une mégacaryopoïèse anormale. Les cultures cellulaires faites à partir des embryons *knock-out* montre une absence de formation de colonies mégacaryocytaires et un développement anormal des globules rouges (Kawada et al. 2001).

Une étude a montré que Fli-1 était impliqué dans l'action synergétique de Gata1 et de Fog-1 pour activer le promoteur du gène spécifique de la lignée mégacaryocytaire GplIB (qui code pour la chaîne de l'intégrine α du récepteur fibrinogène plaquettaire) (Wang et al. 2002). Fli-1 serait également impliqué avec Gata1 dans l'activation de GP1X, GP1b. Il serait aussi impliqué à la fois avec Gata1 et Spl dans l'activation du gène mégacaryocytaire GPVI (Holmes et al. 2002).

1.1.1.6.2.2.2.6 Eklf

Eklf est un facteur de transcription en doigt de zinc (motif de liaison CACC). L'expression de Eklf est restreinte aux lignées érythroïde, mégacaryocytaire et mastocytaire (Perry et al. 2002). Comme tous les membres de la famille Kruppel, EKLf interagit avec GATA1. Plusieurs résultats montrent son implication dans l'érythropoïèse. Des mutations de ce gène provoquent l'apparition de la β -thalassémie chez l'humain et la souris. Le *knock-out* de Eklf résulte en la mort des embryons avec une sévère anémie due à la non-activation du gène de la β -globine adulte. Des sites de liaison pour EKLf sont trouvés sur des éléments régulateurs de plusieurs gènes érythroïdes (Cantor et al. 2002).

1.1.1.6.2.2.2.7 Herf1 (Trim10)

Herf1 a été découvert comme un gène spécifique de la lignée érythroïde. Il semble impliqué dans l'établissement de l'érythropoïèse définitive lors du développement embryonnaire et lors de la différenciation terminale des érythroblastes en érythrocytes. La protéine contient un domaine *RING finger*. L'inhibition de l'expression de Herf1 bloque la différenciation terminale des érythrocytes alors que sa surexpression provoque la maturation de la lignée cellulaire érythroïde MEL (Perry et al. 2002).

1.1.1.6.2.2.2.8. Runx1

Runx1 (ou Aml1) est un facteur de transcription souvent impliqué dans les translocations chromosomiques (AML-1-ETO) des leucémies myéloïdes aiguës (Hitzer

2005, section 1.1.2.4.2.2). RUNX1 se lie à son co-facteur CBF β qui augmente sa capacité à lier l'ADN (sur des motifs TGTGGT).

De récents résultats offrent plusieurs arguments en faveur d'un rôle de Runx1 de façon préférentielle dans la mégacaryopoïèse plutôt que l'érythroïèse. Chez l'humain, RUNX1 est préférentiellement exprimé par les mégacaryoblastes versus les érythroblastes dans lesquels il est absent (Elagib 2003). Il s'exprime très tôt lors de la différenciation avant les marqueurs spécifiques de la lignée mégacaryocytaire. Les souris *knock-out* Runx1^{-/-} meurent en milieu de gestation, dû à un blocage de l'hématopoïèse définitive mais avec une érythroïèse fœtale intacte (Tracey et al. 2000). Un *knock-out* conditionnel de Runx1 a permis de confirmer l'implication de ce gène dans la maturation des mégacaryocytes (Ichikawa et al. 2004).

Cette implication de Runx1 dans la mégacaryopoïèse est confirmée par la présence de nombreux motifs RUNX1 sur les promoteurs de gènes spécifiques de cette lignée (par exemple, les intégrines αIIb et $\beta 3$) (Elagib 2003).

1.1.1.7 Plasticité du système hématopoïétique

De nombreuses observations montrent que le système hématopoïétique est plastique, c'est-à-dire qu'une cellule engagée dans une voie de différenciation a la capacité de changer de voie. Cela se manifeste souvent lors de pathologies comme les leucémies où les cellules peuvent porter les caractéristiques de plusieurs lignées hématopoïétiques différentes (Graf 2002).

Un nouveau concept très complexe émerge actuellement, dans lequel les propriétés d'une cellule 'souche' reflèteraient plutôt *un état* de cette cellule plutôt qu'une entité cellulaire (Zipori 2005, Quesenberry 2006). Ce qui est décrit comme propriété de la cellule souche est la totipotence. Ainsi la cellule (par exemple, HSC) passerait *de façon continue* (concept du 'continuum') d'un état quiescent à un état cyclant (prolifératif). Lors des phases du cycle cellulaire (S, M, G0/G1, G2), la cellule aurait des propriétés différentes: les marqueurs phénotypiques à la surface de la cellule changent, la capacité de repopulation lors d'une greffe également. Ainsi une même cellule aurait toutes les propriétés d'une cellule souche à certaines phases du cycle alors qu'elle aurait un phénotype plus différencié à d'autres moments (figure 1.8, Quesenberry 2006). Ce

concept peut s'appliquer à toutes les cellules progénitrices de la moelle osseuse. Cela explique que c'est parce que la cellule *change d'état de façon continue* qu'elle va être *susceptible de s'engager* dans une voie de différenciation à un moment précis, lorsque le stimulus externe et la phase du cycle sont coordonnés. Ainsi le changement dans l'expression génique peut être favorisé à des moments où la chromatine est dans un état ouvert (euchromatine).

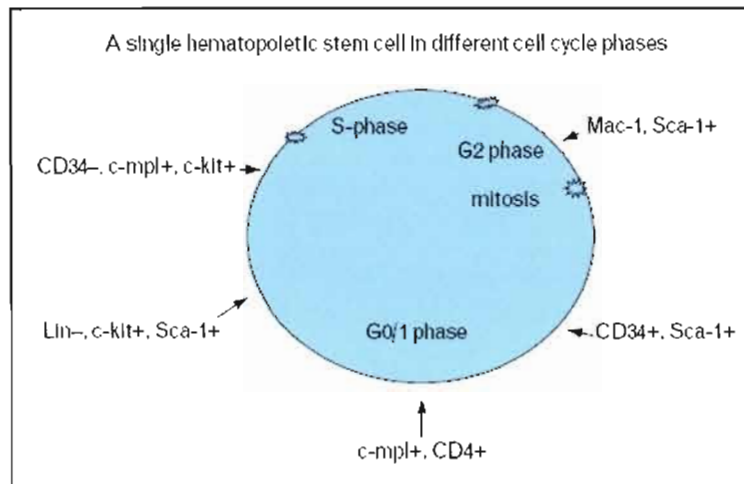


Figure 1.8: Concept du continuum et de l'état 'cellule souche'. Figure tirée de Quesenberry 2006. Cette figure représente une cellule souche durant les différentes phases du cycle cellulaire et les changements phénotypiques qui y sont associés.

Ce concept est important parce qu'il donne une voie pour comprendre comment sont effectuées les décisions d'engagement d'une cellule. Il s'oppose à l'organisation hiérarchique classique du système hématopoïétique (système pyramidal). Dans le concept de hiérarchisation, les cellules s'engagent pas à pas dans la différenciation *d'une manière irréversible*.

Le concept du *continuum* s'élargit sur la notion de dédifférenciation d'une cellule: une cellule qui était dans une voie de différenciation pourrait effectuer des pas en arrière et aller dans une autre voie selon les stimuli reçus (Quesenberry 2006).

Cela explique les nombreux phénomènes de plasticité observés avec les cellules du système hématopoïétique ainsi que la découverte d'intermédiaires de différenciation qui ne se placent pas dans la pyramide classique de l'hématopoïèse (Graf 2002). Un exemple

de dédifférenciation a été observé dans le modèle de leucémie induit par le virus aviaire E26 (*avian acute leukemia virus* E26). Les cellules transformées par le virus sont des érythroblastes qui perdent peu à peu les antigènes spécifiques des cellules érythroïdes pour acquérir des antigènes mégacaryocytaires ainsi que des capacités multipotentes pour devenir des cellules MEP (progéniteur érythromégakaryocytaire). Le virus E26 contient une protéine de fusion cMyb-Ets1. Par comparaison, le virus AEV (*Avian Erythroblastosis Virus*) contient seulement le gène cMyb et induit des cellules transformées portant uniquement des caractéristiques érythroblastiques (McNagny and Graf 2003). Ainsi on voit qu'un facteur de transcription peut guider le phénomène de dédifférenciation.

1.1.1.8 Les désordres de l'hématopoïèse

Une hématopoïèse non parfaitement effectuée a des répercussions sur le fonctionnement de l'organisme. De très nombreuses maladies ou syndromes en proviennent et la discipline médicale associée est nommée hématologie (Corpus médical de la faculté de médecine de Grenoble, www-sante.ujf-grenoble.fr).

Une hématopoïèse défectueuse peut conduire à la production de trop peu de cellules sanguines matures (cytopénies) ou de trop de cellules.

Les cas de cytopénies peuvent provenir par exemple d'une aplasie médullaire, ce qui va provoquer absence ou une trop faible production de tous les types de cellules sanguines. L'aplasie médullaire peut survenir après un traitement de chimiothérapie, après une infection virale (VIH par exemple) ou peut-être idiopathique (enfants immunodéprimés ou enfant bulles). Les cytopénies peuvent provenir également d'un syndrome myélodysplasique correspondant à l'altération des cellules souches de moelle osseuse. La moelle osseuse comprend un trop grand nombre de cellules progénitrices mais dont la différenciation est altérée et il en résulte une pénurie de cellules matures dans le sang périphérique (pénurie de neutrophiles, plaquettes, anémie). Ce syndrome peut précéder le développement d'une leucémie aigue myéloïde.

Il existe aussi des cytopénies correspondant à la trop faible production d'un seul type de cellules sanguines. Il existe par exemple les problèmes d'anémie qui correspondent à une trop faible production de globules rouges et qui conduisent à un mauvais transport de

l'oxygène. Il existe aussi des désordres impliquant une trop faible production de plaquettes conduisant à une mauvaise coagulation du sang.

D'autres désordres peuvent provenir de la production de cellules matures anormales ou non fonctionnelles. Par exemple, la thalassémie ou encore l'anémie falciforme sont des maladies conduisant à une pénurie de globules rouges: résultat d'un défaut de la synthèse des chaînes de la globine dû à une altération génique. L'hémophilie ainsi que la maladie de Willebrand proviennent de défaut dans les facteurs de coagulation sécrétés par les plaquettes.

La maladie de Vasquez ou *polycythemia vera* est au contraire une production trop grande de globules rouges.

Une autre catégorie de désordres hématopoïétiques concerne les hémopathies malignes comme les lymphomes et les leucémies aigues (voir section suivante).

1.1.2. Leucémie

1.1.2.1 Définition

Rodolf Virshow fut un des premiers à décrire la leucémie en 1847. La leucémie est un cancer et elle est définie comme la prolifération non contrôlée ou l'expansion de cellules hématopoïétiques qui n'ont pas la capacité de se différencier normalement en cellules sanguines matures (Sawyers et al, 1991).

La transformation des cellules hématopoïétiques en cellules leucémiques découle de l'acquisition de mutations et de la dérégulation conséquente de certains gènes. Il en résulte une perturbation des étapes de la différenciation. Les cellules leucémiques gardent un état indifférencié et donc le pouvoir de se diviser comme les cellules hématopoïétiques précurseurs. Il résulte donc de l'état de leucémie une pénurie de cellules sanguines matures normales et une prolifération anarchique de cellules leucémiques qui envahissent l'organisme.

Une des caractéristiques de cette maladie est son caractère transplantable. Si on injecte des cellules leucémiques à un animal syngénique (histocompatible), des leucémies peuvent alors être générées dans cet animal. La faculté de transplanter des leucémies

montre que les cellules leucémiques peuvent perdre toute dépendance par rapport aux facteurs de croissance.

Les mutations acquises peuvent être congénitales (par exemple, mutation du gène Ras) ou environnementales (par exemple, exposition au benzène). Un virus nommé HTLV est aussi connu pour provoquer la leucémie (voir section 1.2.4).

La leucémie est une maladie encore en partie incurable (4100 nouveaux cas diagnostiqués au Canada en 2006 et 2200 décès, soit 53% de mortalité) et représente le cancer infantile le plus fréquent au Canada (25,7 % des nouveaux cas de cancer chez les enfants (1997-2001) (Statistiques Canadiennes sur le Cancer, 2006, www.ncic.cancer.ca).

1.1.2.2. Classification des leucémies humaines

La classification des leucémies est très complexe du fait de l'hétérogénéité des leucémies existantes mais a une grande importance car le diagnostic précis des désordres hématopoïétiques détermine le traitement des patients.

Afin de classer ces désordres hématopoïétiques correctement, il faut distinguer les lymphomes des leucémies, et les leucémies chroniques des leucémies aiguës.

Les lymphomes sont des tumeurs malignes et solides du système hématopoïétique alors que les leucémies (tumeurs liquides) se caractérisent par une présence de cellules leucémiques dans tous les organes hématopoïétiques et dans le sang. Un exemple de lymphome est la maladie de Hodgkin qui est un lymphome à prédominance ganglionnaire.

Les leucémies aiguës correspondent à la prolifération rapide des cellules leucémiques. Un traitement immédiat doit être appliqué pour éviter l'envahissement de l'organisme et le décès de l'individu. Elles apparaissent en général chez l'enfant ou chez le jeune adulte.

Les leucémies chroniques sont moins foudroyantes. L'évolution se fait sur des mois ou des années avec des phases de rémission. Elles apparaissent plus généralement chez des personnes âgées.

Les leucémies sont classifiées selon la voie de différenciation et le type de cellules sanguines mis en cause. Ainsi, deux grands groupes de leucémie existent chez l'humain: les leucémies lymphoïdes (atteinte de la lignée hématopoïétique lymphoïde) et les

leucémies myéloïdes (atteinte de la lignée hématopoïétique non-lymphoïde). Ainsi, les 4 grands principaux types de leucémies chez l'humain sont: les leucémies aiguës lymphoïdes (ALLs), les leucémies chroniques lymphoïdes (CLLs), les leucémies aiguës myéloïdes (AML) et les leucémies chroniques myéloïdes (CMLs).

Pour chaque type de leucémies, beaucoup de sous-catégories existent. Les différentes leucémies ont été classifiées dans les années 1970 par une association franco-américano-britannique (classification FAB) et cette classification est basée sur le regroupement des leucémies présentant les mêmes signes cliniques avec une observation de la morphologie des cellules (Bennett et al, 1979, 1985). Elle est encore très utilisée aujourd'hui. Depuis 1995, l'Association Européenne des Pathologistes et la Société pour l'Hématologie ont développé une nouvelle classification nommée classification de l'organisation mondiale de la santé (OMS) (Brunning et al 2001). Cette dernière ne se base plus sur l'observation morphologique mais principalement sur le caryotype (recherche d'anomalies chromosomiques et translocations chromosomiques à valeur de diagnostic et de pronostic) mais aussi sur les signes cliniques et l'immunophénotypage. Depuis, des modifications faites en fusionnant la classification FAB et celle de l'OMS ont été réalisées pour les néoplasmes lymphoïdes (Bennett, 2002) et les néoplasmes myéloïdes (Bennett, 2000).

1.1.2.2.1. Classification des leucémies lymphoïdes

Le tableau 1.1 résume la classification des leucémies aiguës lymphoïdes et des lymphomes (d'après les informations prises de *National Cancer Institute's comprehensive database, Adult and Children Acute Lymphoblastic Leukemia: treatment. PDQ®, health professional versions*).

Tableau 1.1: Classification des néoplasmes humains lymphoïdes

Lignées	Anomalies Génomiques		Pronostic
Lignée B:			bon
	anomalies du caryotype:	hyperdiploidie t(12;21)(p13;q22) t(1;19)(q23;p13) 11q23, t(9;22) t(4;11)(q21;q23)	bon excellent - faible haut risque
Lignée T:			
	gènes impliqués:	MYC, SCL, LYL1, LMO1,-2 HOX11 HOX112	- - favorable faible
	anomalies du caryotype:	t(11.19)(q23;p13) 11q23, t(9;22)	bon faible
Lymphomes non-Hogkiniens			faible
Lymphomes Hogkiniens			

1.1.2.2.2. Classification des leucémies myéloïdes

Le tableau 1.2 résume la classification des leucémies aiguës myéloïdes qui regroupe les classifications de FAB et de l'OMS (*National Cancer Institute's comprehensive database, Adult Acute Myeloid Leukemia: treatment. PDQ®*, health professional version).

Tableau 1.2: Classification des leucémies humaines myéloïdes aiguës

1) AML avec anomalies génomiques caractérisées			Pronostic
	anomalies du caryotype:	t(8;21)(q22;q22) inv(16)(p13q22) t(16;16)(p13q22) t(15;17)(q22;q12) (M3) 11q23; in(3); t(6;9) anomalies chr. 5, ou 7 mutation Flt3 (M0 M1)	bon bon bon bon mauvais mauvais mauvais
2) autres AML			
	myéloblastique:	minimalement différenciée (M0) sans maturation (M1) avec maturation (M2)	
	promyélocytique:	(M3)	bon
	myélomonocytaire:	(M4)	bon
	monoblastique et monocytique:	(M5)	bon
	érythroïde	(M6)	mauvais
	mégakaryoblastique	(M7)	mauvais
	basophilique	-	-

en bleu et entre parenthèses: classification FAB

1.1.2.2.3. Leucémies biphénotypiques

Il existe aussi des leucémies biphénotypiques: les cellules de ces leucémies expriment de façon concomitante des marqueurs spécifiques à la fois de la lignée lymphoïde et de la lignée myéloïde. Ces leucémies sont classées en 2 groupes:

- 1) Leucémies aiguës lymphoïdes (ALLs) positives pour des antigènes myéloïdes
- 2) Leucémies aiguës myéloïdes (AMLs) positives pour des antigènes lymphoïdes.

1.1.2.3. Classification des leucémies murines

La majorité des leucémies obtenues chez la souris le sont par inoculation de rétrovirus (section 1.2) ou par création de souris transgéniques pour une protéine de fusion issue d'une translocation chromosomique.

Chez la souris, plusieurs tentatives de classification ont été publiées durant les cinquante dernières années. En 1954, Dunn publia une première classification systématique des néoplasmes chez la souris (Dunn, 1954). Depuis, les connaissances en hématologie et en hématopoïèse ont progressé et le dernier essai de classification date de 2002: la proposition de Bethesda (Kogan et al. 2002, Morse et al., 2002).

La classification de Bethesda a non seulement pour but d'uniformiser la nomenclature des lésions hématologiques chez la souris mais aussi de faire un parallèle avec les leucémies humaines. Cette classification adapte donc la classification de l'OMS aux leucémies obtenues chez la souris. Les différences qui existent entre les 2 classifications reflètent les différences du point de vue biologique entre la souris et l'humain. Cette classification est basée sur les signes cliniques et morphologiques de la maladie, l'immunophénotypage des cellules leucémiques et les caractéristiques génétiques.

Deux grands groupes de désordres hématopoïétiques sont recensés d'après la classification de Bethesda: le type lymphoïde incluant les désordres associés aux cellules T, B et NK et le type non-lymphoïde incluant tous les autres types ainsi que les leucémies biphénotypiques.

Les tableaux 1.3 et 1.4 illustrent la classification des désordres hématopoïétiques lymphoïdes et non-lymphoïdes établie en 2002.

Tableau 1.3: Désordres hématopoïétiques lymphoïdes chez la souris (Morse et al. 2002)

❖ NEOPLASMES DE TYPE B
Néoplasmes Précurseurs
Néoplasmes Matures
❖ à petites cellules
❖ de la zone marginale splénique
❖ folliculaire
❖ diffus à grandes cellules
❖ lymphome de Burkitt
❖ ressemblant au lymphome de Burkitt
❖ plasmocytaire
❖ des cellules B-NK
❖ NEOPLASMES DE TYPE T
Néoplasmes Précurseurs
Néoplasmes Matures
❖ à petites cellules
❖ des cellules T-NK
Néoplasmes non-caractérisés

Tableau 1.4: Désordres hématopoïétiques non-lymphoïdes chez la souris (Kogan et al. 2002)

❖ LEUCÉMIES NON-LYMPHOIDES
Leucémies Myéloïdes
❖ sans maturation
❖ avec maturation
❖ ressemblant à un désordre
myéloprolifératif
❖ myélomonocytaire
❖ monocytaire
Leucémies Erythroïdes
Leucémies Mégakaryoblastiques
Leucémies Biphénotypique
❖ SARCOMES HÉMATOPOIÉTIQUES
NON_LYMPHOIDES
Sarcomes granulocytaires
Sarcomes Histiocytaires
Sarcomes mastocytaires
❖ DYSPLASIES MYÉLOIDES

1.1.2.4. Le processus de la leucémie

1.1.2.4.1. Généralités

Le processus de la leucémie provient de mutations proches de proto-oncogènes touchant un progéniteur hématopoïétique. La dérégulation de ces gènes fait que la cellule va pouvoir se diviser activement et qu'elle va acquérir également des propriétés d'auto-renouvellement indéfini d'une cellule souche. Classiquement, une cellule a besoin d'un oncogène immortalisant et d'un oncogène impliqué dans la prolifération pour acquérir un phénotype transformé.

Une mutation peut être un évènement déclencheur dans le processus de cancérisation, mais l'étape est progressive, et l'état cancéreux résulte d'une cascade de dérégulation de nombreux proto-oncogènes. Ainsi, de nombreux gènes constituent la signature cancéreuse d'une cellule.

Les translocations chromosomiques sont les anomalies les plus spécifiques des leucémies humaines. Les protéines de fusion résultantes de la translocation chromosomique interfèrent généralement avec la fonction normale d'1 ou des 2 gènes impliqués dans cette fusion.

L'analyse du spectre de mutations impliquées dans les leucémies aiguës humaines montrent que les gènes impliqués peuvent être répartis en 2 groupes complémentaires: ceux qui confèrent un avantage prolifératif et/ou de survie et ceux qui bloquent la différenciation hématopoïétique pour permettre à la cellule de conserver des capacités d'auto-renouvellement. Le premier groupe peut contenir des gènes comme Myc ou Ras alors que le second groupe comprend des gènes impliqués dans l'hématopoïèse normale (Gilliland et al. 2004).

Une cellule leucémique est une cellule bloquée dans le processus de différenciation. Cette cellule est donc plus différenciée que le progéniteur dont elle dérive et moins différenciée que la cellule mature qu'elle aurait dû devenir. La figure 1.9 permet de faire la comparaison entre l'état de différenciation de ces 3 types cellulaires.

Un progéniteur immature exprime des gènes d'indifférenciation alors que la cellule mature n'en exprime plus et inversement pour les gènes de différenciation (figure 1.9 A).

Les gènes qui ne changent pas d'expression entre ces 2 états sont les gènes *housekeeping* (figure 1.9 A). La cellule leucémique comporte moins de gènes d'indifférenciation que le progéniteur (figure 1.9 B) mais plus que la cellule mature (figure 1.9 C). Ces marqueurs d'indifférenciation vont jouer en partie un rôle oncogénique dans la cellule leucémique.

Inversement, la cellule leucémique exprime plus de marqueurs de différenciation que le progéniteur (figure 1.9.B) mais moins que la cellule mature (figure 1.9 C): la cellule leucémique a tout de même été capable d'effectuer quelques étapes de la différenciation.

Parce que la cellule est transformée, elle va exprimer des oncogènes qui sont quiescents à la fois dans le progéniteur et dans la cellule mature, ainsi que des marqueurs aberrants qui sont le reflet de la déviation par rapport à la différenciation normale. De plus, la cellule leucémique exprime moins de gènes suppresseurs de tumeurs que le progéniteur et la cellule mature (figure 1.9 B, C).

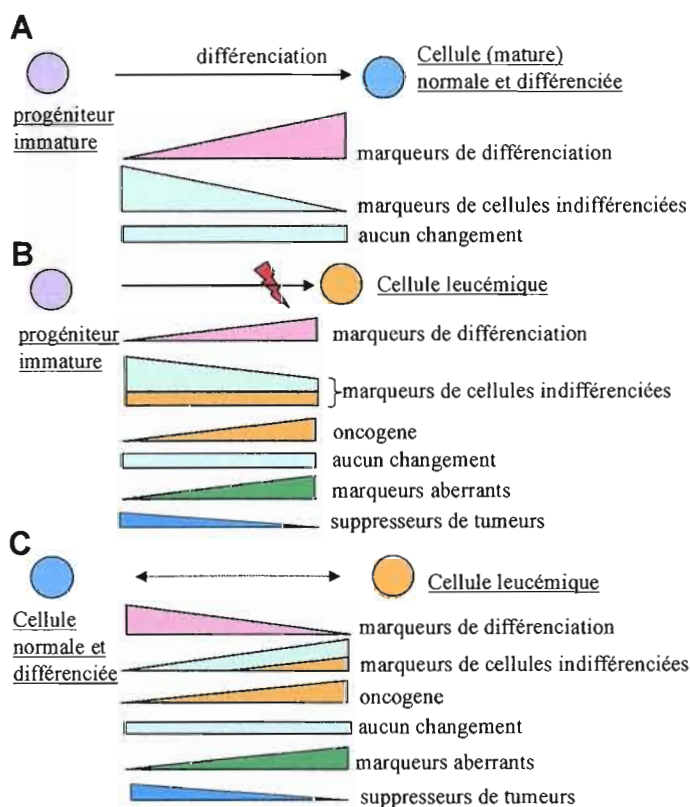


Figure 1.9: Comparaison de l'état de différenciation entre un progéniteur, une cellule mature et une cellule leucémique. (A) Différences entre un progéniteur immature (à gauche) et une cellule mature (à droite). (B) Différences entre un progéniteur

immature (à gauche) et une cellule leucémique (à droite). (C) Différences entre une cellule mature (à gauche) et une cellule leucémique (à droite).

Des exemples de certains gènes et translocations chromosomiques identifiés comme intervenants majeurs dans les leucémies AMLs et ALLs humaines vont être présentés dans la section suivante. Il est à noter que beaucoup de ces gènes sont des déterminants majeurs d'une hématopoïèse non-défectueuse montrant que les 2 éléments sont liés: la leucémie provient d'un déséquilibre de l'hématopoïèse.

1.1.2.4.2. Gènes impliqués dans AML

1.1.2.4.2.1. Flt3

L'activation constitutive de Flt3 (tyrosine kinase) due à des mutations a lieu dans approximativement 30 à 35% des AMLs et dans la majorité des cas, il s'agit de l'acquisition de duplications du gène (*Intern Tandem Duplications: ITDs*). Les mutations de Flt3 (FLT3-ITDs) sont associées à une progression sévère de la maladie et donc à un mauvais pronostic. Ceci rappelle l'importance de connaître les anomalies génétiques des leucémies pour faire le diagnostic mais aussi pour identifier les individus qui ont besoin d'une chimiothérapie plus intensive.

Flt3-ITDs active de nombreuses voies de signalisation impliquées dans la prolifération ou conférant un avantage prolifératif. Les voies de signalisation impliquées sont celles passant par RAS et les MAPKinases, STAT, ou AKT. Une expression forcée de Flt3-ITDs montre que cette forme de Flt3 ne peut pas, à elle seule, induire la leucémie à des souris. Ceci se rapproche des autres tyrosine kinases impliquées dans des translocations chromosomiques comme BCR-ABL, TEL-ABL ou TEL-JAK2 qui ne peuvent pas non plus induire la leucémie à elles seules (Gilliland et al. 2004).

1.1.2.4.2.2. CBF

CBF (*Core Binding Factor*) est un hétérodimère formé de AML1 (RUNX1) et CBF β . Il agit comme un facteur de transcription. Les 2 composants de CBF sont souvent impliqués dans les translocations chromosomiques des leucémies AML (AML1-ETO (t(8;21)), TEL-AML1 (t(12;21)), CBF β -MYH11 (inv16)).

AML1 et CBF β sont essentiels à l'hématopoïèse: il n'y a aucun développement de l'hématopoïèse définitive dans des souris *knock-out* pour AML1 ou CBF β .

Plusieurs études montrent que les protéines de fusion impliquant AML1 ou CBF β sont des versions négatives de CBF et elles vont inhiber plutôt qu'activer les gènes cibles de CBF en recrutant de façon aberrante des co-répresseurs.

Ces protéines de fusion sont importantes dans la pathogenèse des AMLs mais ne peuvent pas à elles seules causer la maladie.

De toutes ces protéines de fusion, AML1-ETO est la plus étudiée. Des études *in vitro* montrent que cette protéine de fusion est capable de conférer un phénotype immortalisant à des cellules transfectées (Gilliland et al. 2004).

1.1.2.4.2.3. PML-RAR α

Les leucémies aiguës promyélocytiques (APL) sont toujours associées à des translocations chromosomiques impliquant le récepteur α de l'acide rétinoïque (RAR). Chacune de ces translocations chromosomiques provoque un blocage de la différenciation des cellules au stade de promyélocytes.

La translocation chromosomique la plus étudiée est PML-RAR α (t(15;17)). C'est une forme négative de RAR α qui inhibe plutôt que stimule les gènes cibles du récepteur par recrutement de co-répresseurs. La protéine de fusion PLM-RAR α seule n'est pas suffisante pour induire la leucémie mais est associée à une inhibition de la différenciation et une augmentation de la capacité à s'auto-renouveler.

L'ATRA (*ALL-trans-retinoic acid*) est un ligand de RAR α et est un agent thérapeutique très efficace pour traiter les APLs surtout en combinaison avec un traitement de chimiothérapie conventionnel. ATRA se lie fortement à la protéine de

fusion, provoquant la dissociation des co-répresseurs et le recrutement des co-activateurs ainsi que la dégradation de la protéine de fusion. Les promyélocytes leucémiques entrent alors de nouveau dans le processus de différenciation et meurent ultimement par apoptose ayant perdu leur capacité d'auto-renouvellement.

Alors que seuls ils n'arrivent pas à induire la leucémie, la combinaison d'un gène muté conférant un avantage prolifératif ou de survie à un gène immortalisant ou d'indifférenciation permet d'induire la leucémie: par exemple l'expression dans des cellules hématopoïétiques murines de Flt3-ITDs et de PML-RAR α , a permis d'obtenir 100 % de pénétrance de la maladie (Gilliland et al. 2004).

1.1.2.4.2.4. GATA1 et AMKL

Chez l'humain, les leucémies mégacaryoblastiques aiguës (AMKL) ont été décrites pour la première fois en 1931 (Von Boros 1931). La leucémie AML-M7 est une forme courante de leucémie infantile, représentant 7 à 10% des leucémies myéloïdes infantiles et est souvent accompagnée de faibles chances de guérison (Athale et al. 2001). Les leucémies AML-M7 se développent également chez l'adulte mais c'est un événement rare représentant environ 1% des AMLs (Pagano et al 2002, Tallman et al. 2000).

Les AMKLs sont cependant particulièrement fréquentes chez les enfants souffrant du syndrome de Down (trisomie 21). Même si ces individus souffrent plus fréquemment de T-ALLs, ils ont un risque 500 fois plus élevé que la population normale de développer une AMKL (DS-AMK). Mais contrairement aux leucémies AMKLs non liées au syndrome de Down, les DS-AMKLs répondent plus favorablement au traitement de chimiothérapie (Hitzler 2005).

De façon intéressante, de nombreux groupes de recherche ont trouvé une mutation dans le gène Gata1 (situé sur le chromosome X) spécifiquement dans les cellules d'individus souffrant de DS-AMKLs. Cette mutation est absente dans le cas des AMKLs non liées au syndrome de Down ainsi que dans les cas de T-ALLs liées au syndrome de Down. De plus, les patients en rémission ayant souffert de DS-AMKL ne présentent plus de mutation dans le génome des cellules de la moelle osseuse. Ces résultats indiquent que la perte du GATA-1 sauvage constitue une étape clé dans la pathogenèse des DS-AMKLs (Hitzler 2005).

Les mutations de Gata1 sont détectées dès la naissance dans les individus souffrant du syndrome de Down, indiquant qu'ils représentent un événement se passant *in utero*. Les mutations sont situées dans l'exon2 et il en résulte une protéine GATA1 tronquée du côté N-terminal et qui ne possède pas de domaine d'activation (Hitzler 2005).

1.1.2.4.3. Gènes impliqués dans ALL

1.1.2.4.3.1. B- ALLs

La translocation chromosomique t(12;21)(p13;q22) conduisant à la protéine de fusion TEL-AML1 est la translocation la plus fréquente pour les leucémies infantiles ALL de type B (25 % des leucémies infantiles ALLs). La présence de cette translocation est associée à d'excellentes chances de survie (90% de guérison) (Armstrong et Graf, 2005).

La protéine de fusion contient le domaine basique hélice-boucle-hélice de TEL fusionné avec le domaine de liaison à l'ADN et de transactivation de AML1. La fonction de AML1 a déjà été expliquée précédemment notamment dans la translocation AML1-ETO. TEL, comme AML1, a aussi une implication dans l'hématopoïèse et un *knock-out* conditionnel de TEL montre qu'il est indispensable à l'hématopoïèse définitive. La protéine de fusion résulterait en une perte de fonction de TEL et de AML1, ce qui causerait la leucémie.

Une autre translocation très fréquente dans les B-ALLs est la translocation t(9;22), produisant la protéine de fusion BCR-ABL (chromosome de Philadelphie). Contrairement à TEL-AML1, elle est beaucoup plus fréquente chez les adultes (33% des B-ALLs adultes) que chez les enfants (5% des B-ALLs infantiles). Elle est associée à de très faibles chances de guérison et il faut recourir à la transplantation de la moelle osseuse pour espérer une guérison (Armstrong et Graf, 2005).

1.1.2.4.3.2. T- ALLs

Beaucoup de facteurs de transcription comme MYC, SCL (TAL1) ou LYL, et qui ne sont pas normalement exprimés dans la lignée T sont des sites préférentiels de translocations chromosomiques des T-ALLs. Il est à noter également que beaucoup de

gènes sont réarrangés en étant fusionnés à proximité de la chaîne β du TCR (notamment les gènes LMO1, LMO2, ou HOX11).

L'expression aberrante de SCL par la translocation t(1;14) ou par une délétion intragénique est présente dans 25% des leucémies infantiles T-ALLs. C'est d'ailleurs dans ce contexte que ce gène a été découvert. SCL est connu pour former un complexe pentamérique avec E2A, LMO2, GATA1 et LDB1. L'expression ectopique de SCL dans les cellules T pourrait activer des gènes normalement quiescents et c'est ainsi qu'il jouerait son rôle dans la leucémogénèse. Mais une hypothèse plus réaliste consiste à penser que SCL aurait comme rôle d'inactiver E2A en le séquestrant dans les complexes pentamériques. En effet, des observations allant dans le même sens que cette hypothèse montrent que des souris déficientes pour E2A développent des T-ALLs. De plus, SCL n'a besoin que de son domaine de liaison à E2A et non celui de liaison à l'ADN pour induire les T-ALLs (Armstrong et Graf, 2005).

Chez l'adulte, l'implication du gène homéotique HOX11 intervient dans 33% des cas alors qu'il intervient dans 5% des cas seulement chez l'enfant.

1.1.2.4.3.3. MLL

Le dernier gène présenté en exemple dans cette section est MLL, dû à son implication dans un grand nombre de translocations chromosomiques.

MLL est un gène qui intervient dans les leucémies du nourrisson et qui est associé à un très mauvais pronostic. Il n'est pas spécifique des ALLs à cet âge car plus de 70% des nourrissons avec une leucémie aiguë ont une translocation chromosomique 11q23 impliquant MLL. Ces leucémies sont soit de type AML soit de type ALL et souvent biphénotypiques, d'où le nom donné à ce gène: *Mixed Lineage Leukemias gene* (MLL). Une quarantaine de translocations chromosomiques impliquant MLL ont été découvertes.

L'analyse du *knock-out* de Mll suggère qu'il joue un rôle important dans le développement et dans l'hématopoïèse et son rôle est de réguler l'expression des gènes Hox. Ainsi le rôle de MLL dans la leucémogénèse semblerait être relié à la régulation des gènes HOX car de nombreuses études montrent l'implication de ces gènes dans la leucémie (par exemple translocation t(7;11)(p15;p15) conduisant à la protéine HOXA9-NUP98 dans les AMLs) (Armstrong et Graf, 2005).

Beaucoup de gènes impliqués dans les leucémies ont été trouvés par identification des anomalies du caryotype des cellules de patients leucémiques. Les régions comprenant les translocations chromosomiques ont alors été clonées et les gènes ont été caractérisés. Des souris ont alors été créées avec un transgène comportant les protéines de fusion. Ainsi ces souris transgéniques représentent les modèles majoritairement utilisés pour étudier la leucémie. Dans une section suivante (section 1.2), un autre modèle va être décrit: il s'agit de l'utilisation de rétrovirus murins de la leucémie. Ce modèle présente le grand avantage de permettre également l'identification d'oncogènes qui ne sont pas impliqués dans les translocations chromosomiques.

1.1.2.4.4. Concept de cellules souches leucémiques

Un concept récent et pourtant bien établi est la présence dans la population leucémique de cellules moins différenciées portant beaucoup de caractéristiques des cellules souches hématopoïétiques. Ces cellules sont nommées cellules souches leucémiques (LSC: *leukemia stem cells*). Les LSCs ont maintenu ou réacquis leur capacité de prolifération indéfinie à cause de l'acquisition de mutations ou de changements épigénétiques (Passegue and Weissmann, 2005). Les LSCs donnent naissance à la population leucémique et sont responsables de la récurrence de la maladie (Zou 2007). Elles ont beaucoup de caractéristiques communes avec les HSCs mais elles se différencient tout de même par une capacité plus grande d'auto-renouvellement et elles portent également des marqueurs phénotypiques différents.

La question se posant actuellement est: quelles sont les cellules de la moelle osseuse, HSCs ou progéniteurs plus différenciés, qui se transforment en LSC? Comme les LSCs ont des caractéristiques très proches des HSCs, la transformation des HSCs en LSCs semble la plus probable surtout si l'on pense à un système hématopoïétique de type pyramidal. Cependant, les résultats d'expériences montrent que des progéniteurs hématopoïétiques peuvent être transformés et peuvent donner naissance à des populations leucémiques. Ainsi, ces résultats suggèrent que certaines mutations permettent de générer des LSCs à partir de progéniteurs plus différenciés que des HSCs. Les mutations permettraient aux cellules progénitrices de réacquérir premièrement la capacité de s'auto-renouveler (Gilliland et al. 2004). Ceci est tout en fait en accord avec le concept du 'continuum' s'opposant au système pyramidal (section 1.1.1.7). On pourrait alors

supposer que LSC est un état plutôt qu'une entité cellulaire, et que dans la population leucémique générée et suivant des signaux extra-cellulaires, certaines cellules leucémiques passeraient tour à tour et continuellement d'un état plus différencié au stade de LSCs, permettant ainsi la conservation du pool leucémique. Ainsi pour une approche thérapeutique efficace, il faudrait détruire toute la population leucémique simultanément avec toute la population de LSCs.

1.1.2.4.5. Thérapies

La leucémie a la caractéristique d'être un cancer très hétérogène. De nombreux sous-types existent. Ces sous-types peuvent être définis par les anomalies génétiques décrites précédemment ou par la réponse au traitement. La réponse au traitement dépend non seulement du type de leucémie mais aussi des variations génétiques individuelles notamment dans les protéines impliquées dans le transport des métabolites exogènes. En tenant compte de ces variables, l'approche la plus efficace consiste à adapter l'intensité du traitement à chaque patient grâce à un diagnostic extrêmement précis de sa condition (Yeoh et al. 2002). Les agents anti-cancer sont généralement toxiques et donnés à des doses à la limite de la toxicité. En estimant précisément l'intensité du traitement suivant la prédiction de la réponse au traitement et l'agressivité de la leucémie, le patient risque de recevoir des doses insuffisantes ou exagérées.

L'établissement du profil génique des cellules leucémiques grâce à des micropuces à haute densité a fait énormément avancer le diagnostic. La première étude de profilage génique performante fut réalisée par Golub *et al.* (1999) qui ont montré la possibilité de faire la différence entre des patients AML et des patients ALL grâce à la signature moléculaire. L'observation des données relatives à la survie des patients dont provenaient les échantillons a montré que les différents sous-types étaient corrélés à de différents pronostics. De nombreuses autres études de profilage génique ont alors été réalisées sur les différents sous-types de leucémie connus.

La corrélation entre les anomalies génétiques et les chances de survie apportent beaucoup au diagnostic. Par exemple, les leucémies avec la protéine de fusion E2A-PBX1 répondent très faiblement à la thérapie conventionnelle mais sont guéries à 80% quand elles sont traitées avec des thérapies plus intensives. Les leucémies avec la protéine de fusion BCR-ABL répondent très mal à la chimiothérapie conventionnelle et de

meilleurs résultats sont obtenus avec une transplantation de moelle osseuse. Des études de profilage génique ont pu montrer des différences d'expression des gènes entre des patients qui rechutent et ceux qui guérissent dans un même sous-type de leucémie.

Le profil génique permet de faire un lien plus solide entre les lésions génétiques et les chances de survie. De nouveaux sous-types ne correspondant pas à une lésion génétique déterminée ont également été découverts. La figure 1.10 montre un très bel exemple de profilage génique réalisé sur des échantillons de ALLs (Yeoh et al. 2002). Les échantillons présentant les mêmes lésions géniques (E2A-PBX1, MLL, BCR-ABL) se regroupent ensemble pour former une signature génique commune. Le groupe *Novel* indique un sous-type qui ne correspondait à aucune lésion génétique prédéfinie.

Le profilage génique semble assez perfectionné maintenant pour remplacer les méthodes actuelles qui nécessitent un immunophénotypage, une analyse du caryotype ainsi que des diagnostics moléculaires.

Les nombreux sous-types identifiés pour chaque classe de leucémie mettent en évidence que la transformation des cellules peut se produire par la dérégulation de plusieurs voie de signalisation (Yeoh et al. 2002).

Un autre point important est la grande variation individuelle de la réponse au traitement. Pour un même sous-type de leucémie, les patients vont réagir différemment parce qu'ils portent des variations génétiques dans leurs gènes codant par exemple pour les enzymes responsables de métaboliser les substances exogènes ou les transporteurs de ces substances. Cela influence l'efficacité et la toxicité du traitement. Un exemple de différence individuelle due à un polymorphisme génétique implique le gène TMPT (Thiopurine méthyltransférase). L'enzyme codée permet la méthylation des thiopurines utilisées pour traiter les ALLs. Certaines personnes ont des allèles produisant une TMPT très active alors que d'autres portent un allèle déficient. Des tests génétiques ont été établis pour déterminer quels allèles sont portés par un individu. Si un individu porte des allèles productifs, des doses fortes de thiopurines doivent être administrées pour que le traitement soit efficace. Si les allèles portés sont peu efficaces, le traitement doit être réduit en raison de la forte toxicité (Cheok and Evans 2006).

Le profilage génique permet également d'identifier les groupes de gènes responsables de la résistance au traitement mais l'analyse des allèles responsables de la résistance est

aussi importante pour apporter un traitement individualisé. Il existe par exemple des micropuces spécialisées dans la détection des SNPs (*Single Nucleotide Polymorphisms*).

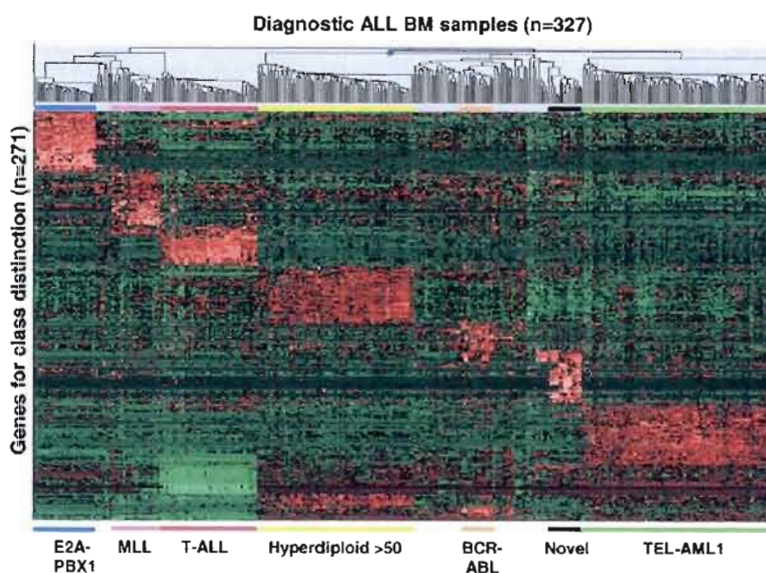


Figure 1.10: Cluster hiérarchique de 327 échantillons de leucémies ALL (colonnes) et 271 gènes (lignes). Figure tirée de Yeoh et *al.* 2002. En rouge: gènes fortement exprimés, en vert: gènes faiblement exprimés.

Des stratégies moléculaires ciblées se développent de plus en plus pour traiter les leucémies. Elles offrent des thérapies avec moins d'effets secondaires et de toxicité. Elles se développent à partir de la connaissance plus grande des gènes et des voies de signalisation impliqués dans les différents sous-types de leucémie. Elles consistent à trouver des inhibiteurs d'un gène cible ou de voies de signalisation particulières. On peut citer comme un exemple de réussite de stratégies moléculaires la combinaison de ATRA avec des thérapies conventionnelles ciblant la protéine de fusion PML-RAR α (Gilliland et al. 2004), l'inhibiteur imatinib (Gleevec) de la tyrosine kinase ABL ciblant la protéine de fusion BCR-ABL (Druker et al, 2001). Les mutations activatrices de FLT3 sont corrélées avec un très mauvais pronostic peu importe le sous-type de leucémies. Ainsi de nombreux chercheurs tentent d'établir un inhibiteur spécifique de cette tyrosine kinase. De nombreux inhibiteurs ont été développés à l'heure actuelle et les essais sont

encourageants ; car ils induisent l'apoptose des cellules exprimant FLT3 (Gilliland et al. 2004)

Une autre stratégie de thérapie vise les cellules souches de la leucémie. Ces cellules qui seraient quiescentes répondent moins aux traitements classiques de chimiothérapie. Détruire ces cellules pourrait certainement diminuer le problème de rechute fréquemment rencontré lors du traitement des leucémies et autres cancers. Une récente étude montre la reconnaissance spécifique des cellules souches leucémique AML humaines transplantées à des souris grâce à l'antigène de surface CD44. Le blocage de cet antigène de surface permet d'éliminer ces cellules souches leucémiques dans ces souris et diminuent la capacité de repopulation de la leucémie et ouvre ainsi la voie à de possibles thérapies (Jin et al. 2006).

1.2. Les rétrovirus murins de la leucémie

1.2.1. Découverte des rétrovirus murins de la leucémie

Les virus murins de la leucémie (MuLVs pour *Murine Leukemia Viruses*) sont les agents responsables de la transmission de la leucémie d'une souris à l'autre. Les MuLVs ont été découverts au début du siècle dernier à une période pendant laquelle les scientifiques ont commencé à élever des souris afin de les utiliser en expérimentation. L'observation des différentes souches de souris consanguines montrait que certaines étaient plus susceptibles de développer des tumeurs ou des leucémies que d'autres. La probabilité d'apparition des tumeurs ne suivait pas exactement les lois de Mendel, faisant supposer qu'un facteur autre que génétique était responsable. Par exemple, la souche AK fut développée par Furth en 1933 et elle permit la découverte du premier MuLV par Gross en 1951. Gross fut aussi le premier à obtenir un virus assez virulent pouvant induire des leucémies lorsque inoculé à des souriceaux nouveaux-nés. Il s'agit du virus nommé 'Passage A'. Gross a aussi découvert le virus nommé Akv (Gross 1951, 1957).

Plusieurs autres chercheurs ont alors réussi à la même époque à isoler différents virus à partir de souris de laboratoire. Ils ont souvent donné leurs noms aux virus: virus de Friend, (Friend, 1957), Moloney, (Moloney, 1950), Raucher (Rauscher 1962), Graffi (Graffi, 1957). On peut citer également l'isolement à la même époque de RadLV (Kaplan,

1964). Le fait que les souches de souris de laboratoire ont été établies par seulement quelques chercheurs dans le monde a pour conséquence que les différents isolats de MuLVs proviennent de quelques mêmes ancêtres. Leurs différences proviennent de recombinaisons acquises lors de passages *in vivo*.

Les chercheurs ont ensuite voulu savoir si ces virus étaient présents à l'état naturel et ont donc porté leur attention sur les souris sauvages. Effectivement, l'ADN génomique des souris sauvages contient la trace d'infections ancestrales sous la forme de fragments proviraux et dans de rares cas de provirus entiers. Parmi les rétrovirus non-défectifs qui ont été trouvés chez les souris sauvages, on peut citer la découverte en Californie des virus dits de classe amphotropique et du virus Cas-Br-E (Gardner 1978). Des virus défectifs endogènes sont également présents dans les souches de souris de laboratoires puisqu'elles dérivent elles-mêmes de souris sauvages.

C'est à partir des années 1960 que les virus murins de la leucémie ont commencé à être utilisés comme outils de recherche sur le cancer aux Etats-Unis (*Virus Cancer program, National Cancer Institute*) et une grande avancée dans la compréhension de ces virus a été faite avec la découverte de l'enzyme rétro-transcriptase par D. Baltimore et H.M Temin en 1970.

Chronologie de la découverte des MuLVs (Stephenson, 1980):

1910-1935 : L'ère des virus de la leucémie commence par la découverte du virus du sarcome du poulet (Rous) comme agent transmissible et responsable de la maladie.

1936-1950 : Cette époque est marquée par l'étude sur les MMTVs (*mouse mammary tumor virus*) et sur le virus du papillome. Ils découvrent à cette époque la période de latence entre la transmission de l'agent infectieux et l'apparition des tumeurs.

1951-1970 : C'est la période de découverte de plusieurs virus murins de la leucémie qui sont encore utilisés de nos jours comme modèle d'étude de la pathogenèse rétrovirale ou du cancer. La progression de la compréhension de la pathogenèse rétrovirale est due à l'amélioration des techniques de biologie cellulaire et moléculaire. L'intégration du génome du virus dans celui de la cellule hôte est découverte.

1980-2006 : augmentation de la compréhension du cycle rétroviral et pathogenèse associée, et découverte de nombreux oncogènes grâce aux MuLVs. Utilisation des

rétrovirus comme outil de recherche sur le cancer et développement des thérapies géniques.

1.2.2 Structure des MuLVs

Les informations nécessaires à l'écriture des sections 1.2.2, 1.2.3 et 1.2.4 ont été prises majoritairement dans le livre *Retroviruses* (Coffin et al. 1997)

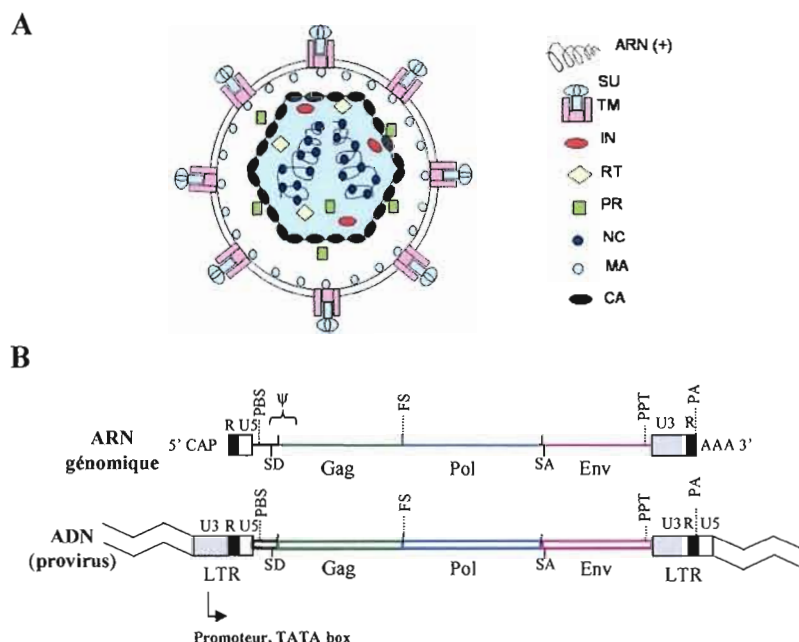
Les MuLVs sont des *rétrovirus* à ARN et sont classifiés comme rétrovirus de mammifères de type C. Récemment, la classification des rétrovirus a été redéfinie (www.ncbi.nlm.nih.gov): les MuLVs sont dans le groupe VI des virus à ARN rétro-transcrit, dans la famille des Retroviridae, et sont du genre gammarétrovirus.

Les MuLVs sont classifiés selon leur tropisme (type cellulaire et souches de souris qu'ils sont capables d'infecter) et selon qu'ils sont endogènes (présents dans la lignée germinale d'une souche de souris) ou exogènes (provenant d'isolats mais non présents dans la lignée germinale).

Ce sont des virus à enveloppe formée d'une bicouche lipidique issue de la membrane cytoplasmique des cellules hôtes. Ces virus ont comme matériel génétique deux molécules d'acides ribonucléiques simple brin de polarité positive qui codent pour des protéines structurales et fonctionnelles.

La particule virale contient tous les éléments nécessaires à la réplication du virus à l'intérieur d'une cellule hôte et est faite de façon à ce qu'il soit bien protégé lors de son passage dans l'environnement extracellulaire. Le schéma d'une particule virale est présenté sur la figure 1.11 A. Le génome du virion se trouve protégé au cœur de la particule et est complexé avec des protéines structurales pour former la nucléocapside. Une coque interne, la capsid, est formée autour de la nucléocapside. A ce niveau-là se trouvent également des protéines fonctionnelles: l'intégrase, la rétro-transcriptase et la protéase. La capsid est entourée par une coque externe supplémentaire, la matrice, qui est elle-même en association étroite avec la bicouche lipidique de l'enveloppe. La bicouche lipidique est parsemée de protéines de surface (SU et TM) accrochées à la matrice et qui vont reconnaître des récepteurs membranaires à la surface des cellules hôtes.

Les gènes codant pour les protéines structurales et fonctionnelles du virus sont organisés en 3 régions principales, légèrement chevauchantes, le long du génome viral. Ces 3 régions correspondent aux gènes Gag, Pol et Env qui codent pour des protéines précurseurs qui vont être clivées par la protéase pour former des protéines matures (figure 1.11 B). Gag code pour les protéines structurales de la matrice (MA), capsid (CA) et



nucléocapside (NC). Pol code pour des protéines fonctionnelles: protéase (PR), RNase H, rétro-transcriptase (RT) et intégrase (IN). L'enveloppe code pour les protéines de surface (SU) et transmembranaire (TM).

Figure 1.11: Organisation des virions et des génomes viraux, schémas adaptés de *Retroviruses* (Coffin et al. 1997). (A) Schéma simplifié d'un virion. SU: protéine de surface; TM: protéine transmembranaire; IN: intégrase; RT: rétro-transcriptase; PR: protéase; NC: nucléocapside; MA: matrice; CA: capsid. (B) Organisation génomique. PBS: primer binding site; SD: site donneur de splicing; SA: site accepteur d'épissage; FS: frameshift; PPT: polypurine tract; PA: queue polyA; LTR: Long terminal repeat.

Aux extrémités 5' et 3' du génome viral se trouvent des structures dont l'organisation va légèrement différer s'il s'agit du génome viral (dans la particule) ou du provirus (forme intégrée dans l'ADN de la cellule hôte). Lorsque l'on parle du provirus, chaque extrémité identique à l'autre est appelée LTR (*Long Terminal Repeat*). (figure 1.11 B) Les LTRs sont composés des régions U3, R (*Repeat region*) et U5 et jouent un rôle essentiel dans le

déroulement de la transcription inverse de l'ARN génomique en ADN proviral ainsi que dans la transcription des ARN messagers à partir du provirus. La région U3 est un fort promoteur liant des facteurs de transcription présents dans les cellules hôtes. Lors de la réaction de rétro-transcription, les brins d'ARN génomique vont se replier sur eux-mêmes formant une boucle ce qui permet la duplication en 5' de U3 initialement en position 3' et celle de U5 en 3'.

1.2.3 Cycle de réplication rétrovirale

Dans l'organisme infecté, le virus va obligatoirement suivre un cycle de réplication pour se multiplier (figure 1.12). Le virus entre dans la cellule hôte grâce à la reconnaissance d'un récepteur membranaire par les protéines de surface de l'enveloppe. Il rentre dans la cellule par endocytose et va commencer son cycle de réplication.

Un virus comporte le minimum de protéines et d'information génétique nécessaires à sa multiplication. Afin de se propager, il va utiliser à ses fins la machinerie de la cellule dans laquelle il vient de pénétrer.

En utilisant sa propre enzyme, la rétro-transcriptase (ou transcriptase inverse), le virus rétro-transcrit son génome qui est sous forme d'ARN en molécule d'ADN double-brin. La molécule d'ADN double-brin va migrer dans le noyau et s'intégrer, grâce à l'intégrase, dans le génome de la cellule cible. L'ADN viral intégré, appelé alors provirus, va être considéré par la cellule hôte comme de l'ADN génomique et plus particulièrement comme un gène. C'est ainsi que le virus détourne la machinerie cellulaire pour produire beaucoup de nouvelles molécules d'ARN. Certaines de ces molécules d'ARN viral vont être destinées à intégrer les nouvelles particules virales qui vont se former, et d'autres vont être traduites en protéines. Les protéines synthétisées sont celles codées par les 3 gènes, Gag, Pol et Env (voir section 1.2.2). L'assemblage des nouvelles particules s'effectue et celles-ci sortent de la cellule par bourgeonnement, emportant avec elles la bicouche lipidique de la cellule hôte transformée en enveloppe. Une particule entrée dans une cellule permet la production de nombreuses autres particules virales qui vont sortir de la cellule. Ce phénomène d'amplification ainsi que l'utilisation de la machinerie cellulaire sont des preuves de l'adaptation ingénieuse des virus dans le monde vivant.

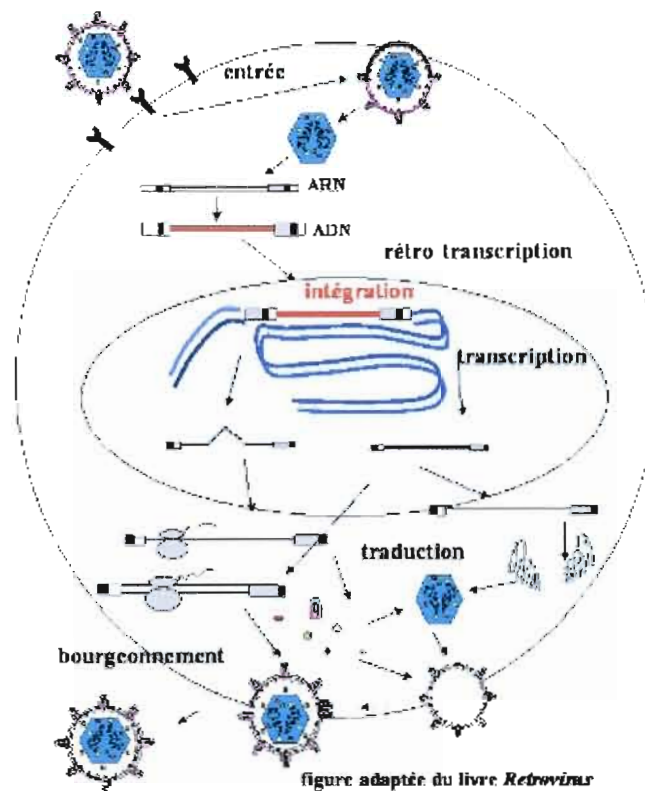


Figure 1.12: Cycle de réplication rétrovirale. Les principales étapes du cycle sont montrées. Le cercle correspond au dessin de la cellule hôte; l'ovale correspond au noyau. Le schéma est adapté de *Retrovirus*, (Coffin et al. 1997).

1.2.4 Induction de la leucémie

La capacité des MuLVs à induire la leucémie provient de *l'intégration de leur génome* dans celui de la cellule hôte.

Certains rétrovirus sont capables d'induire la leucémie ou le cancer parce que leur génome a *capturé* par transduction un oncogène dans son propre génome (rétrovirus défectifs). On peut citer le MuLV C58-MSV qui contient l'oncogène *v-ras* (Fredrickson et al., 1987) ou le MuLV Ab-MLV contenant l'oncogène *Abl* (Abelson and Rabstein,

1970) Ainsi, lorsqu'ils infectent une cellule hôte, l'oncogène porté par le génome viral devient partie intégrante de l'ADN cellulaire et de nombreuses copies du gène sont transcrites. Les protéines oncogéniques correspondantes dérèglent la prolifération de la cellule qui devient cancéreuse.

Les autres MuLVs, qui n'ont pas la particularité de contenir un oncogène, et qui ont donc un génome entier tel que décrit dans la section 1.2.2, ont également la capacité d'induire la leucémie à cause de leur intégration dans le génome cellulaire. Trois mécanismes généraux sont décrits: 1) mutagenèse insertionnelle par insertion de promoteur ou de enhancer; 2) production de protéines tronquées; 3) inhibition de suppresseurs de tumeur. La figure 1.13 schématise le principe de la mutagenèse insertionnelle ainsi que celui de la production de protéines tronquées.

- 1) La mutagenèse insertionnelle provoque la sur-expression d'un proto-oncogène situé à proximité de l'intégration rétrovirale. La transcription peut se faire à partir du promoteur cellulaire du proto-oncogène et la transcription est favorisée par la présence d'éléments *enhancers* des LTRs viraux situés en amont ou en aval du gène (cas noté 1 sur la figure 1.13). Dans ce cas, la double-hélice d'ADN formant des structures tertiaires rapprochant les éléments, le virus peut agir sans être intégré à proximité du proto-oncogène, en amont ou en aval de celui-ci. Le deuxième cas correspond à l'utilisation du promoteur viral pour initier la transcription du proto-oncogène (cas noté 2 sur la figure 1.13). Dans ce cas, le provirus doit être intégré à proximité du gène, en amont et dans une orientation sens.
- 2) La production de protéines tronquées est provoquée par l'intégration du virus à l'intérieur de la région transcriptionnelle du proto-oncogène. La troncation a pour conséquence de produire une protéine dont les fonctions oncogéniques sont accrues.
- 3) Le dernier principe correspond à l'inhibition de la transcription de gènes suppresseurs de tumeur ou la production d'une protéine inactive par troncation, cela par intégration rétrovirale. La protéine normalement codée par les gènes suppresseurs de tumeur ne peut alors plus agir sur l'inhibition de la prolifération cellulaire ou l'induction de l'apoptose.

L'ADN viral va s'intégrer dans le génome de la cellule hôte de façon aléatoire même si

des sites préférentiels sont connus. Pour les MuLVs, ses sites ont d'abord été rapportés correspondant à des structures chromatiniennees ouvertes c'est-à-dire sensibles à la ADNase I. ou proches de sites d'initiation de la transcription et d'îlots CpG (Panet and Cedar, 1977, Lewinski et al. 2006). Ces sites semblent un peu différents entre les classes de rétrovirus et par exemple pour VIH et le génome humain, les sites préférentiels se trouvent dans les unités transcriptionnelles actives (Lewinski et al. 2006). Il semble que l'intégrase est le facteur déterminant de la spécificité de chaque classe de rétrovirus (Lewinski et al. 2006).

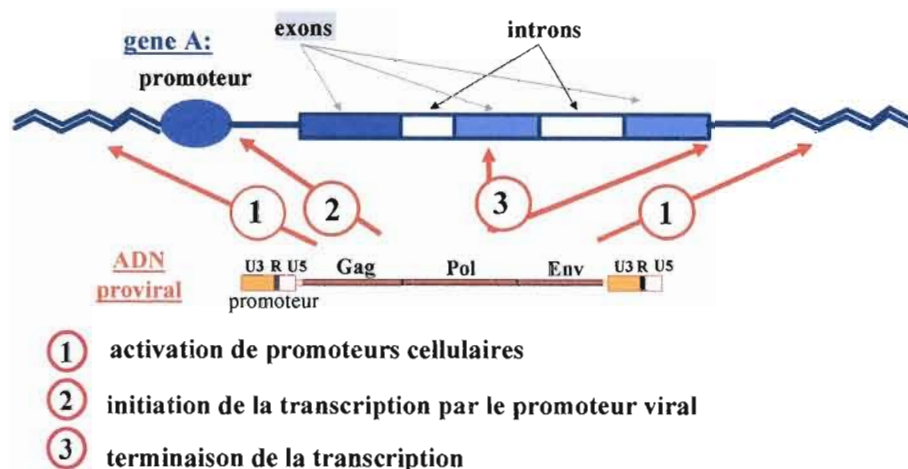


Figure 1.13: Mécanismes de la mutagenèse insertionnelle. Le schéma est adapté de *Retroviruses*, (Coffin et al. 1997).

Si le virus arrive à s'intégrer à proximité d'un proto-oncogène et à déréguler son expression, cela confère à la cellule un avantage prolifératif. La cellule se divise activement et donne naissance à des cellules portant la même intégration. Les cellules deviennent cancéreuses et forment une population clonale de cellules leucémiques. C'est au début des années 1980, que les intégrations rétrovirales ont commencé à être mises en évidence à proximité de proto-oncogènes ainsi que la dérégulation d'intégration de ces proto-oncogènes dans les tumeurs (Hayward et al. 1981). Parmi les oncogènes, c-myc est très fréquemment ciblé par les intégrations rétrovirales. Par exemple, de multiples intégrations à proximité de c-myc ont été trouvées avec les rétrovirus de Friend (Dreyfus et al. 1990), de Graffi (Denicourt et al. 1999), ou le Gross Passage A (Villemur et al. 1987)

Entre le moment de l'inoculation du virus à des souris et l'apparition de la leucémie, il existe toujours un temps de latence assez grand même s'il diffère suivant les types de MuLVs. Ce temps de latence est nécessaire pour qu'une intégration rétrovirale cible et dérégule un proto-oncogène. Ainsi les cellules cibles d'une souris infectée peuvent également contenir des intégrations rétrovirales qui n'ont aucune relation avec le développement de la maladie.

Un *site commun d'intégration rétrovirale* (CIS) peut être mis en évidence grâce à des outils de biologie moléculaire. Il s'agit de l'intégration d'un virus dans une même région chromosomique dans des tumeurs issues de différentes souris. Les CIS sont trouvés en général, à proximité, ou au sein de proto-oncogènes, qui sont les premiers intervenants dans le déclenchement de la leucémie, comme c-myc par exemple. Ainsi la recherche de sites communs d'intégration rétrovirale dans les tumeurs induites est un outil puissant pour découvrir de nouveaux proto-oncogènes et c'est ainsi que les MuLVs ont participé grandement à la compréhension du cancer dans les dernières années. Cette recherche est facilitée actuellement grâce à l'obtention de la séquence du génome de la souris.

Pourquoi les MuLVs induisent-ils spécifiquement des leucémies et non toutes les autres formes de cancer? La spécificité des MuLVs à induire la leucémie provient du fait qu'ils ciblent les progéniteurs hématopoïétiques. Des études ont montré que dans les premières heures après l'inoculation virale, les virus se trouvent dans les ostéoclastes (et les ostéoblastes) de la moelle osseuse puis ensuite les particules virales infectent les progéniteurs hématopoïétiques (Okimoto and Fan, 1999). Afin de pénétrer dans leurs cellules cibles, les rétrovirus utilisent des récepteurs spécifiques des cellules. Par exemple, les MuLVs écotropiques utilisent le récepteur mCAT1 (*murine cationic amino acid transporter 1*) (Albritton et al. 1989). Les MuLVs ont également développé un LTR bien adapté à la réplication dans les cellules hématopoïétiques, puisqu'il peut lier des facteurs de transcription fortement exprimés dans ce type de cellules (Golemis et al. 1989, Barat and Rassart, 1998).

Les MuLVs offrent l'avantage d'induire des leucémies qui sont de très bons modèles pour étudier les leucémies humaines et les gènes qui ont été découverts grâce aux MuLVs ont également un rôle oncogénique chez l'humain (exemple: c-myc).

Il existe un seul rétrovirus chez l'humain capable d'induire la leucémie. Il est nommé HTLV-I (*Human T- leukemia/lymphotropic virus type I*) (Poiesz et al. 1980). Il est associé à l'apparition de désordres lymphoprolifératifs et dans 1% des cas, des leucémies de type T se développent avec une latence très longue (20 ans). HTLV-1 provoque d'autres maladies comme des syndromes neurologiques. Il est particulièrement fréquent au Japon, en Afrique, en Amérique du sud et dans le bassin des Caraïbes (Feuer and Green, 2005). HTLV-II a également été trouvé chez un patient atteint d'une leucémie lymphoïde (*hairy cell leukemia*) mais ce variant est moins pathogénique que HTLV-I (Feuer and Green, 2005). Récemment un troisième variant a été identifié au Cameroun, HTLV-III mais la pathologie à laquelle il est associé n'est pas connue (Calattini et al. 2005).

Il est à noter que des rétrovirus autres que les MuLVs peuvent induire d'autres pathologies comme par exemple le BM5eco qui induit un syndrome immunitaire proche du SIDA ou des variants neuropathologiques des MuLVs (exemple de PVC-211, variant neuropathologique du virus de Friend, (Hoffman et al. 1992)).

1.2.5. Le rétrovirus Graffi

1.2.5.1. Les travaux d'Arnold Graffi

Arnold Graffi a découvert dans les années 1950 un virus capable d'induire la leucémie chez les souris et lui a donné son nom (figure 1.14, Graffi, 1957). Il était à cette époque professeur à l'Institut pour la Recherche sur le Cancer de l'Académie Allemande des Sciences (Berlin-Buch). A. Graffi est décédé à l'âge de 96 ans en 2006 (Wunderlich and Bielka, 2006).

A cette époque, l'isolement des virus se faisait en récoltant l'extrait acellulaire à partir d'organes infiltrés de souris souffrant de leucémie. Les techniques de purification se sont développées peu à peu (techniques de centrifugation et de gradients de densité) afin d'obtenir un extrait d'une pureté plus grande (comportant moins de cellules ou de débris cellulaires) et d'une concentration plus grande en particules virales. Cependant, les chercheurs travaillaient souvent à l'époque avec des mélanges de virus.

Arnold Graffi travaillait avec 3 souches principales de souris, nommées XVII, Agnès-Bluhm et CBA. Ces souches ont disparu à l'heure actuelle excepté la souche CBA. Le rétrovirus Graffi a été isolé à partir de l'extrait acellulaire de tumeurs d'Ehrlich nommées *Sarcoma I* et *Sarcoma II* (Graffi 1957).

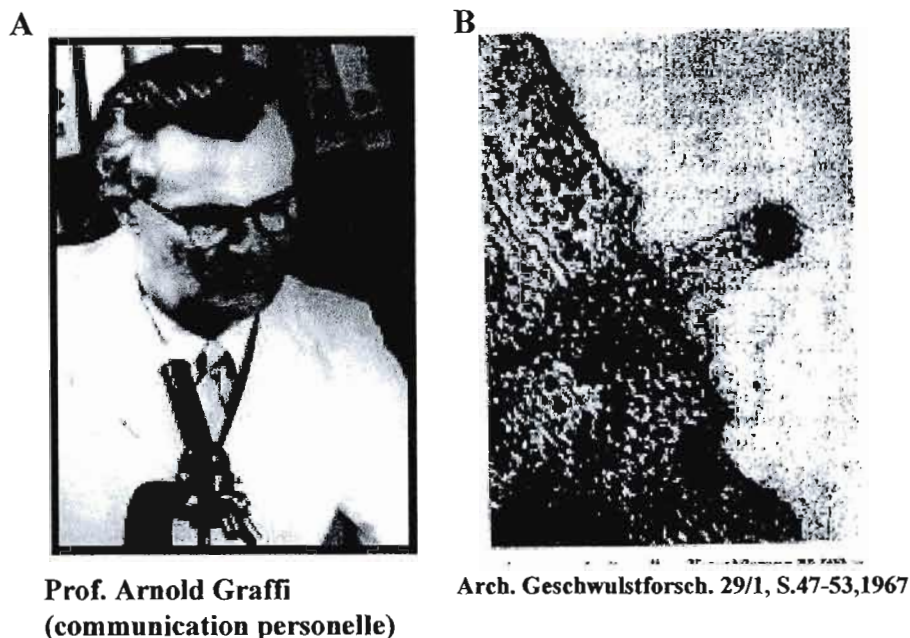


Figure 1.14: Arnold Graffi et le virus Graffi. (A). Photographie d'Arnold Graffi. (B) Photographie en microscopie électronique d'une particule du virus Graffi bourgeonnant hors d'une cellule.

Lors des premières injections des extraits acellulaires, les leucémies induites étaient presque exclusivement des leucémies de type myéloïde. L'extrait acellulaire était peu virulent, avec une latence d'induction très longue (31 semaines) et seulement un faible pourcentage des souris développait la maladie. 70% des souris développaient des chloroleucémies, un type particulier de leucémie myéloïde qui a pour caractéristique une coloration verdâtre des tissus hématopoïétiques due à une abondance de myéloperoxidase (Graffi 1957).

Au cours des différents passages, l'extrait acellulaire présentait une virulence accrue (temps de latence plus court et plus grande incidence d'apparition des leucémies) mais en même temps apparut également le phénomène de la diversification hématologique. Ce

phénomène correspond au fait que le pourcentage de chloroleucémies induites diminue et qu'il y a apparition d'autres types de leucémies même si elle sont classées majoritairement comme myéloïdes (Fey and Graffi, 1965a, 1965b, 1968, Graffi et al., 1966). La première fois qu'ils ont constaté la diversification hématologique, c'est quand ils ont injecté à des souris un extrait acellulaire provenant de rats inoculés auparavant. Puis ils ont constaté la diversification hématologique de façon généralisée et n'ont jamais réussi à reproduire un pourcentage aussi élevé de chloroleucémies qu'originellement. D'autres chercheurs travaillaient aussi sur le virus Graffi: ils trouvaient que certaines leucémies induites étaient de type lymphoïde (Fiore-Donati and Chieco-Bianchi 1966).

Ainsi, les différents types de leucémies répertoriées dans ces articles selon la classification de l'époque étaient les suivantes : 1) Leucémies myéloïdes différenciées, 2) Leucémies myéloïdes immatures, 3) Chloroleucémies, 4) Leucémies paramyéloblastiques, 5) Leucémies myéloïde-réticulaires, 6) Leucémies réticulaires, 7) Leucémies réticulaire-lymphatiques, 8) Leucémies lymphatiques, 9) Leucémies érythroblastiques.

A partir de l'observation de la diversification hématologique, ils ont commencé à faire des expériences pour en comprendre le mécanisme (sans succès), ce qui fut résumé dans les publications citées ci-dessus. (Fey and Graffi, 1965a, 1965b, 1968, Graffi et al., 1966). Ce qu'ils constatent d'après les résultats de leurs expériences, c'est que le type de leucémie induit majoritairement dépend essentiellement de l'extrait acellulaire utilisé au départ. Ainsi, s'ils prennent une leucémie de type chloroleucémie au départ et qu'ils en injectent l'extrait acellulaire à d'autres souris, ils obtiennent un très grand pourcentage de leucémies de type chloroleucémie ou myéloïde et s'ils prennent une leucémie de type lymphoïde, ils obtiennent beaucoup de leucémies de type lymphoïde. Mais il existe toujours dans tous les cas une diversification hématologique et elle s'accroît lors des passages successifs. Ils ont aussi constaté que le pourcentage de chloroleucémies diminue avec la diminution du temps de latence. Ils font alors une corrélation avec l'âge de la souris et la maturation du système hématopoïétique.

Ils ont trois hypothèses pour expliquer la diversification hématologique : 1) il y a un mélange dans l'extrait acellulaire et chaque virus du mélange est responsable d'un type de leucémies; 2) Il existe seulement un type de virus, et il infecte au hasard un type de cellule hématopoïétique et transforme ces cellules en cellules cancéreuses; 3) Le virus se transforme selon les conditions et la transformation résulte en une différence de

tropisme. Leur hypothèse favorite est la 3^{ème}. Ils pensent que le virus lorsqu'il bourgeonne de la cellule hôte dans laquelle il s'est répliqué, emporte avec lui un peu de l'enveloppe cellulaire. Cette enveloppe cellulaire porterait des récepteurs qui seraient responsables de l'histotropisme.

Leur caractérisation était basée sur une observation morphologique des frottis sanguins et certaines réactions histochimiques. Mais à cette époque les connaissances sur les cellules du système immunitaire étaient plus réduites et il est donc difficile de faire un parallèle entre les résultats d'hier et d'aujourd'hui. Par exemple que sont réellement les leucémies réticulaires? Enfin, fait important: on sait également aujourd'hui qu'ils travaillaient avec un mélange viral.

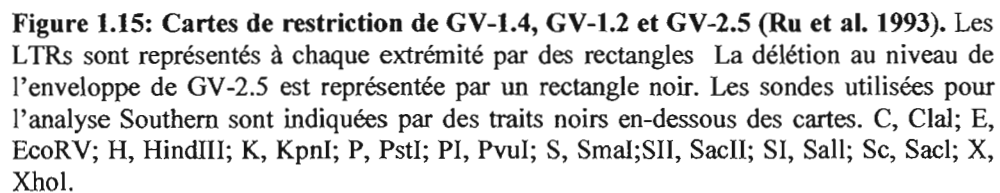
1.2.5.2. Clonage du rétrovirus Graffi

En 1993, le laboratoire d'Eric Rassart a réussi à cloner le génome de 2 variants du rétrovirus Graffi (Ru et al.1993). Le clonage procure l'avantage de pouvoir produire un virus pur et de permettre diverses manipulations et expérimentations.

Les variants ont été isolés à partir de cellules chroniquement infectées par l'extrait acellulaire original du virus Graffi (données par Nathalie Teich). Afin de réaliser le clonage, l'ADN de haut poids moléculaire a été isolé à partir des cellules chroniquement infectées et une banque de clones a été réalisée. Plusieurs dizaines de clones ont été obtenus à l'aide d'une sonde spécifique provenant de la région U3 du LTR du rétrovirus Moloney. Cette sonde permettait de reconnaître spécifiquement les rétrovirus écotropiques exogènes. A l'issue de cette analyse, 3 types de clones comportant apparemment le génome rétroviral entier avec 2 LTRs ont été sélectionnés pour poursuivre l'étude. Les 3 clones furent nommés GV-1.2, GV-1.4, et GV-2.5.

Des analyses Southern Blot plus poussées ont été réalisées et ont montré que GV-2.5 avait une délétion dans la région de l'enveloppe. De plus, l'absence d'une région d'environ 60bp dans la région U3 de GV-1.4 et présente dans celle de GV-1.2 a été mise en évidence. La figure 1.15 représente la carte de restriction enzymatique de ces 3 variants.

Les clones GV-1.2 et GV-1.4 ont permis la production de virus infectieux alors que GV-2.5 probablement défectif ne donna pas de virus infectieux.



La séquence nucléotidique des LTRs de GV-1.2 et GV-1.4 a été élucidée (numéros d'accèsion L14415 pour GV-1.2 et L14416 pour GV-1.4) et cela a permis de voir que les LTRs de GV-1.2 et GV-1.4 étaient identiques excepté une répétition de 60 pb présente dans le U3 du GV-1.2 et absente dans celui de GV-1.4.

Malgré la classification myéloïde, la présence de caractéristiques lymphoïdes des tumeurs est largement discutée dans l'article. Les caractéristiques lymphoïdes correspondent à un réarrangement du TCR β et de la chaîne lourde de l'immunoglobuline dans certaines tumeurs. Ce phénomène a été appelé 'infidélité de lignage', caractérisant les présences concomitantes de caractères myéloïdes et lymphoïdes. La première hypothèse pour expliquer le phénomène était que le virus Graffi infectait un progéniteur déjà engagé dans la lignée lymphoïde mais qui aurait provoqué une altération de la différenciation lymphoïde et une réactivation des caractères myéloïdes. La deuxième hypothèse proposée correspond à l'inverse, c'est-à-dire l'infection d'un progéniteur myéloïde dont le programme lymphoïde aurait été réactivé.

1.2.5.3. Travaux sur la région U3 du virus Graffi

Partant de la séquence nucléotidique des LTRs de GV-1.2 et GV-1.4, 2 études ont été effectuées pour caractériser la région U3 du virus Graffi (Barat et Rassart, 1998 a et b). Comme le virus semblait avoir la particularité unique parmi les autres MuLVs d'induire des leucémies myéloïdes, le but était de trouver quels facteurs de transcription pouvaient se lier spécifiquement au U3 du virus Graffi.

La région entière de U3 a été utilisée pour cette analyse et non pas seulement la région *enhancer* comme c'est le cas dans de nombreuses études. Le *enhancer* est une région bien spécifique de U3 situé en amont de la boîte CAAT et de la boîte TATA et qui lie de nombreux facteurs de transcription jouant donc le rôle d'activateur de la transcription. C'est dans cette région que se trouve la répétition de GV-1.2. La région *enhancer* a été notamment bien étudiée pour les virus de Moloney et de Friend et de nombreux sites consensus liant des facteurs de transcription ont été identifiés (Golemis et al, 1989, Golemis and Speck 1990, Manley et al. 1989).

La première étude a tenté d'élucider les sites liant des facteurs de transcription du U3 de GV-1.4 et celui du virus Cas-Br-E en émettant l'hypothèse qu'ils n'induisaient pas les mêmes pathologies (Barat and Rassart 1998a). Les expériences ont été effectuées avec des lignées cellulaires de diverses origines hématopoïétiques. De nombreux sites liant les facteurs de transcription ont été identifiés et ils se superposent en majorité avec ceux déjà identifiés pour les virus de Moloney et de Friend (éléments CORE, LVb,...). Le facteur CBF β a été identifié comme celui se fixant sur l'élément CORE. De plus, un site en

dehors de la région *enhancer* semble différencier le U3 de GV-1.4 et celui de Cas-Br-E. Dans une seconde étude (Barat and Rassart 1998b), ces travaux ont été continués et il a été montré que les facteurs GATA-1, GATA-2, et GATA-3 se fixaient et trans-activaient les U3 de GV-1.4 et Cas-Br-E.

L'activité de transcription de la région U3 de GV-1.2 et GV-1.4 a été testée dans plusieurs lignées cellulaires (myéloïde, érythroïde, lymphoïdes, fibroblastique) (Barat and Rassart 1998a, figure 1.16). Les activités de transcription ont été testées en comparaison avec différents virus (Cas-Br-E, Friend, Moloney, VL3V13, SV40). Le résultat montre que GV-1.2 et GV-1.4 ont une forte activité de transcription dans toutes les lignées testées. GV-1.4 a une activité plus faible que GV-1.2 qui montre le taux d'activité le plus élevé par rapport aux autres virus testés. GV-1.2 a même une activité de transcription plus forte que le virus Moloney connu pour induire uniquement des lymphomes de type T dans les lignées lymphoïdes. Ces résultats étaient inattendus car on pensait voir une activité de transcription des variants GV-1.2 et GV-1.4 plus forte ou même restreinte à la lignée myéloïde testée.

TABLE 1
Relative Transcriptional Activity of Retroviral LTRs in Various Cell Lines

	Erythroid	Myeloid		T Lymphoid		Cos
	D1B	M1	32Dcl3	BW5147	Jurkat	
Cas-Br-E	65	43	36	19	5	43
Friend	100 ^a	88	70	35	8	57
GV1.2	90	100 ^a	100 ^a	61	100 ^a	100 ^a
GV1.4	81	88	80	39	80	86
Moloney	43	39	55	100 ^a	80	23
VL3V13	63	62	56	46	38	49
SV40	3	5	4	0	1	84

Note. Reporter plasmids containing U3 region of various MuLVs upstream of the CAT gene were transfected in the indicated cell lines. Results are the mean value of at least three experiments. The standard deviation was generally not above 10%.

^a The highest CAT activity in each cell line was given the arbitrary value of 100, and the activity of other LTR-CAT constructs are expressed as percentage activity of this highest LTR.

Figure 1.16: Activité de transcription de la région U3 de plusieurs virus (Barat et Rassart, 1998 a).

Il est à noter que les LTRs des rétrovirus sont formidablement adaptés. Par exemple, le rétrovirus Graffi fixe des facteurs de transcription qui sont non seulement fortement exprimés dans les cellules dans lesquelles les particules virales se multiplient (GATA-1, -2 -3) mais aussi qui sont responsables de l'état leucémique (CORE: CBF β et AML1).

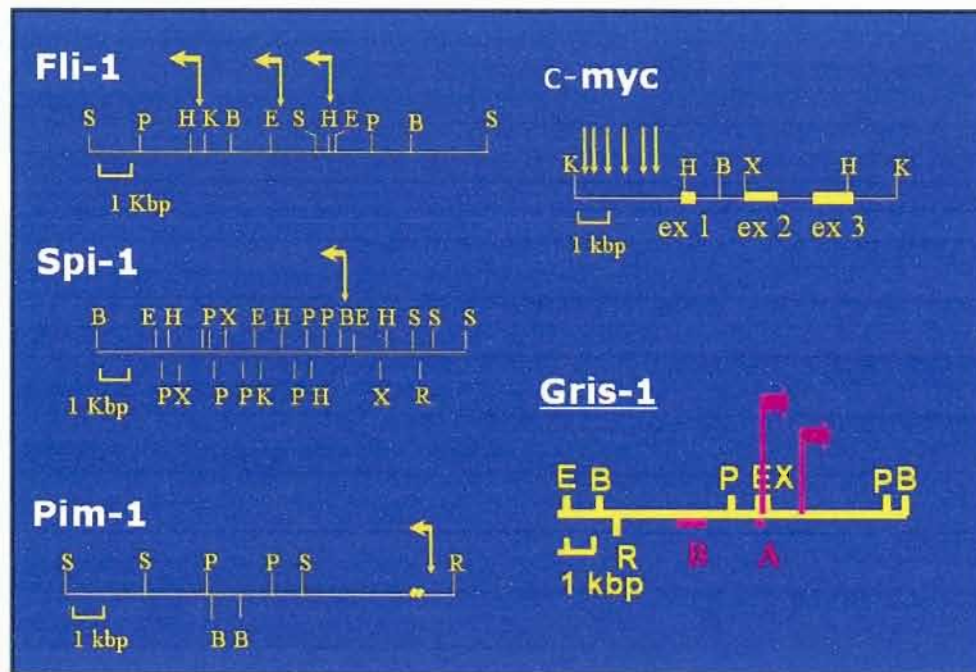
1.2.5.4. Recherche de sites communs d'intégrations

Afin d'utiliser le rétrovirus Graffi comme un modèle pour étudier le cancer, les intégrations à proximité d'oncogènes connus pour être ciblés par différents MuLVs ont été recherchées (Denicourt et al., 1999). L'hypothèse était de connaître quels étaient les oncogènes spécifiques des leucémies myéloïdes. Des intégrations ont été retrouvées dans c-myc (6 tumeurs, 20%), Fli-1 (3 tumeurs, 10%), Spi-1/Pu.1 (1 tumeur, 3,3 %) et Pim 1 (1 tumeur, 3,3 %) (figure 1.17). Les tumeurs réarrangées pour c-myc ne montraient pas de sur-expression par rapport aux autres car toutes les tumeurs montraient une expression très forte de ce gène. Il a été souligné dans l'article que c-myc est effectivement exprimé de façon abondante dans toutes les cellules proliférantes et n'est régulé négativement que lors de la différenciation. Pim1 a été trouvé dans une tumeur montrant déjà un réarrangement pour c-myc. Les intégrations dans Fli-1 ont montré une sur-expression spécifique dans les tumeurs réarrangées.

La recherche de nouveaux sites communs d'intégration (C.I.S) dans les tumeurs induites par le rétrovirus Graffi a conduit à la découverte de Gris-1 (*Graffi integration site 1*) (Denicourt et al. 2003). Le site est localisé sur le chromosome 6 de la souris et les intégrations rétrovirales sont trouvées dans 13% des tumeurs analysées (induites avec GV-1.4). A l'époque de la découverte de ce C.I.S, le génome de la souris n'était pas encore public et plusieurs expérimentations ont dû être effectuées pour découvrir les gènes associés au C.I.S. Il s'agit de la cycline D2 située à 85kbp des intégrations. Le virus intégré provoque la surexpression du gène de la cycline D2 ainsi que d'une isoforme tronquée de ce gène issu d'un épissage alternatif jusque-là non caractérisée. Une étude poussée de l'isoforme tronqué montre qu'il a des propriétés oncogéniques (Denicourt et al., 2007, *Oncogene in Press*).

L'équipe de R.Elwel et I.Touv a effectué une recherche à grande échelle des sites d'intégrations rétrovirales de GV-1.4 dans des souris FVB/N par une technique de PCR inverse (Erkeland et al. 2004). Ils ont analysé 69 tumeurs et trouvé 94 sites d'intégrations différents. Parmi les 94 intégrations, 79 représentaient des CIS dont environ 1/3 étaient des sites d'insertions connus. Seulement 19 insertions ont été testées pour confirmation et

l'expression des gènes à proximité des gènes n'a également pas été analysée. Un site commun qui a été plus profondément analysé par cette équipe correspond au ciblage du



gène YY1 par GV-1.4 (Erkeland et al. 2003).

Figure 1.17: Intégrations rétrovirales trouvées dans des tumeurs induites par GV-1.4 (Denicourt et al. 1999). La région génomique ciblée est schématisée par un trait. La carte de restriction enzymatique est indiquée. Les intégrations virales ainsi que l'orientation du virus sont indiquées grâce à des flèches.

CHAPITRE II

CARACTÉRISATION DES LEUCÉMIES INDUITES PAR LE RETROVIRUS GRAFFI

2.1. Avant-propos

Le but principal du projet de thèse présenté dans ce manuscrit est d'utiliser le rétrovirus murin Graffi pour étudier le phénomène de la leucémie. La première étape essentielle fut d'établir le modèle d'étude, c'est-à-dire d'avoir une très bonne connaissance des leucémies induites par le rétrovirus murin Graffi.

Le projet a ainsi commencé par une étude approfondie de la pathogenèse associée au rétrovirus Graffi. Cette nouvelle caractérisation était nécessaire au vu de la complexité des résultats obtenus antérieurement. Premièrement, Ru *et al.* (Ru et al. 1993) bien que classifiant les leucémies comme myéloïde discutait grandement de la complexité des résultats et du phénomène d'infidélité de lignage: les leucémies portaient également des caractéristiques lymphoïdes (réarrangement des gènes du TCRb et Jh) (Ru et al. 1993, section 1.2.5.2). De plus, des travaux postérieurs à l'article de 1993 (Ru et al. 1993) et antérieurs à 2001, mettaient en évidence des leucémies de type T chez des souris Balb/c inoculées avec GV-1.2 (travaux de C. Barat).

Ainsi, la complexité de la pathogenèse associée au rétrovirus Graffi ayant pu être suspectée, une nouvelle approche de caractérisation a été effectuée en ne se basant pas sur les résultats antérieurs. Pour cela, les 2 variants rétroviraux, GV-1.2 et GV-1.4 ont été inoculés à 3 souches de souris, Balb/c, NFS et FVB/n. Les souris ont été observées pour détecter l'apparition des signes pathologiques. Elles ont été sacrifiées au stade terminal de la maladie et les organes hématopoïétiques ont été utilisés pour effectuer différentes analyses. Cela a apporté de nombreux nouveaux éléments sur la pathogenèse associée au rétrovirus Graffi et cela a fait l'objet d'un article: *Novel insights into the pathogenesis of*

the Graffi murine leukemia retrovirus, Voisin V, Barat C, Hoang T, Rassart E, J Virol. 2006 Apr;80(8):4026-37. L'auteur C. Barat a participé dans cet article à la caractérisation des leucémies (leucémies induites par inoculation de GV-1.2 dans les souris Balb/c) et a corrigé le manuscrit. L'auteure T. Hoang a fourni l'expertise en immunophénotypage des leucémies et a aidé à corriger le manuscrit.

2.2. Article: 'Novel insights into the pathogenesis of the Graffi Murine Leukemia Retrovirus'

2.2.1. Abstract

The Graffi murine leukemia virus (MuLV) was isolated in 1954 by Arnold Graffi, who characterized it as a myeloid leukemia-inducing retrovirus. He and his team, however, soon observed the intriguing phenomenon of hematological diversification, which corresponded to a decrease of myeloid leukemias and an increase of other types of leukemias. Recently, we derived two different molecular clones corresponding to ecotropic nondefective genomes that were named GV-1.2 and GV-1.4. The induced leukemias were classified as myeloid based on morphological analysis of blood smears. In this study, we further characterized the two variants of the Graffi murine retrovirus, GV-1.2 and GV-1.4, in three different strains of mice. We show that the Graffi MuLV is a multipotent retrovirus capable of inducing both lymphoid (T- and B-cell) and nonlymphoid (myeloid, erythroid, megakaryocytic) leukemia. Many of these are very complex with concomitant expression of different hematopoietic lineages. Interestingly, a high percentage of megakaryocytic leukemias, a type of leukemia rarely observed with MuLVs, arise in the FVB/n strain of mice. The genetic backgrounds of the different strains of mice influence greatly the results. Furthermore, the enhancer region, different for GV-1.2 and GV-1.4, plays a pivotal role in the disease specificity: GV-1.2 induces more lymphoid leukemias, and GV-1.4 induces more non lymphoid ones.

2.2.2. Introduction

About 50 years ago Arnold Graffi, at the Cancer Research Institute in Berlin, Germany, isolated a new retrovirus capable of inducing leukemia in a specific strain of mice, named Agnes-Bluhm, bred in his laboratory. Graffi (Graffi, 1957) induced a large proportion of myeloid leukemias with a very high incidence of chloroleukemias (70%), characterized by a greenish coloration of the lymph nodes, and he classified the virus as a myeloid leukemia-inducing retrovirus.

Working on better characterizing the pathogenesis of the virus, Graffi and his team were soon confronted with complex results. They observed the intriguing phenomenon of hematological diversification, which corresponded to a decrease in the percentage of chloroleukemias and an increase in other types of leukemias (Fey, 1969, Fey and Graffi, 1965, Graffi, et al. 1966). This virus could induce multiple kinds of myeloid leukemias, from immature to differentiated forms, but also reticular, lymphoid, and erythroid leukemias and mixed forms (based on the classification at this time).

It is very likely that Graffi was working at that time with a viral mixture. In 1993, Ru et al. (Ru, et al. 1993) cloned two nondefective ecotropic retroviral genomes from NIH-3T3 cell lines chronically infected with the original Graffi acellular extract (a gift of Nathalie Teich). These two viral variants were called GV-1.2 and GV-1.4. The genomic restriction map and the sequences of their long terminal repeats (LTRs) (GenBank accession numbers L14415 and L14416 for GV-1.2 and GV-1.4, respectively) showed that they differ in the U3 region, one 60-bp segment being duplicated in GV-1.2 and not in GV-1.4. GV-1.2 induces disease with a shorter latency period. Based mainly on morphological analysis of blood smears, the leukemias induced with GV-1.2 and GV-1.4 were described as mainly myeloid/granulocytic when injected into newborn BALB/c and NFS mice (Ru, et al. 1993). In several tumors, we observed myeloid characteristics (myeloperoxidase [MPO] staining) together with rearrangements of the germinal configuration of the T-cell receptor β (TCR β) gene and/or of the immunoglobulin heavy chain gene, specific for T-cell and B-cell lineages, respectively. This phenomenon was called lineage infidelity.

In this report, we extended our characterization of the Graffi retrovirus. GV-1.2 and GV-1.4 were inoculated into three strains of mice (BALB/c, NFS, and FVB/n), and more

than 100 leukemias were analyzed by flow cytometry with several different hematopoietic markers combined with molecular biology tools.

The complete genome sequences of the two variants were determined and used to complement the immunophenotyping analysis. This study reveals that the Graffi murine leukemia virus (MuLV) is a complex virus capable to induce a large spectrum of leukemias. It replicates efficiently in several cell types, as already suggested in studies of the LTR (Barat and Rassart. 1998 a,b). It is multipotent and induces both lymphoid and nonlymphoid leukemia, including megakaryocytic leukemias, which are quite rare in MuLV-induced pathologies.

2.2.3. Materials and Methods

2.2.3.1. Viruses, inoculation of mice, and tissue collection

Viral stocks of GV-1.2 and GV-1.4 were made from cell culture supernatants derived from chronically infected NIH-3T3 cells. The culture supernatant of the infected cells was collected, centrifuged, aliquoted, and kept at -80°C . Newborn (<24 h-old) mice of strains BALB/c, NFS, and FVB/n were inoculated intraperitoneally with 0.1 ml of filtered virus. The viral titers were 3.10^6 PFU for GV-1.2 and 1.10^6 PFU for GV-1.4. The mice were checked routinely for clinical signs of disease (loss of weight, anemia, enlarged spleen). Moribund mice were sacrificed, and several organs (thymus, lymph nodes, spleen, liver, kidneys, brain) were frozen on dry ice and stored at -80°C . A sterile cell suspension from the spleen, thymus, and lymph nodes was prepared by gently teasing the organs apart in cold RPMI cell culture medium and by passing them through nylon mesh. Viable cells were used for flow cytometry analyses or stored at -80°C in 90% fetal calf serum and 10% dimethyl sulfoxide. Bone marrow was recovered by flushing cold RPMI cell culture medium through the femur with a 26G3/8 needle. Blood and spleen samples were incubated in hypotonic ammonium chloride solution to eliminate the erythrocytes.

2.2.3.2. Immunophenotyping.

10^6 cells were suspended in 100 μ l of standard phosphate-buffered saline buffer. For detection of surface antigens, the antibodies (quantities calculated from a titration experiment) were added, and the cells were incubated for 30 min in the dark. The cells were then washed with cold phosphate-buffered saline and resuspended in 500 μ l of a standard isotonic solution (Hematal; Fisher). The antibodies (BD Pharmingen, Mississauga, Canada) used were as follows: fluorescein isothiocyanate (FITC)-conjugated rat anti-mouse CD4, FITC anti-CD3, phycoerythrin (PE)-conjugated anti-CD90, FITC anti-CD45R/B220, FITC anti-CD11b (Mac-1, M1/70), FITC anti-CD71 (transferrin receptor), FITC anti-CD41(glycoprotein IIB), FITC anti-CD34, FITC anti-CD117 (cKit, stem cell factor ligand), FITC anti-Sca1, PE anti-CD8a, PE anti-Gr-1 (Ly6G/Ly6C, RB6-8C5), and PE anti-TER119 (glycophorin A). Isotype controls (BD Pharmingen) were included for each antibody used. Flow cytometric analyses were done with a fluorescence-activated cell sorter scanner (FACScan; Becton Dickinson), and the data were processed with the WinMDI software. CD3 together with CD90, CD4 together with CD8, Mac-1 together with Gr-1, and Ter119 together with CD71 were used in double-staining combinations.

2.2.3.3. RNA extraction and Northern blotting

Total RNA samples were extracted from the tissues using the TRIzol reagent kit (Invitrogen, Burlington, Canada) and stored at -80°C . Fifteen micrograms of total RNA samples were denatured with formamide and separated on a 1% agarose gel containing formaldehyde. The RNA was transferred to a nylon membrane, UV fixed, and hybridized with DNA probes labeled by the random primer extension method using oligohexamers (GE Healthcare, Baie d'Urfé, Canada), and approximately 2×10^6 cpm/ml was used. The hybridization process was done as described previously (Sambrook and D. Russell. 2001). The cDNA fragments used as probes were as follows: myeloperoxidase (S. Gisselbrecht), PU.1 (F. Moreau-Gachelin), EpoR (H. Lodish), Fli-1 (A. Bernstein), FOG-1 (S. Orkin), c-mpl, GpIIb, and PF4 (F. Morlé). The expression of beta-actin was used as a control for equal RNA loading and to normalize relative expression data.

2.2.3.4. DNA extraction and Southern blotting.

The genomic DNA was extracted from normal and leukemic tissues by standard procedures as described previously (Sambrook and D. Russell. 2001). The DNA was digested with appropriate restriction enzymes, separated on a 0.8% agarose gel, and transferred to nylon membranes (Ambion) for hybridization. A 600-bp TCR β probe (kindly provided by H. Fan) and a 600-bp probe corresponding to the Jh region (provided by F. Alt) were used to detect T-cell receptor and heavy chain immunoglobulin gene rearrangements, respectively.

2.2.3.5. PCR amplification of the proviral LTR.

PCR amplification of the proviral LTR region was performed by using genomic tumor DNAs with the forward primer 5'-CCCCACCATAAGGCTTAGCAAGCTAG-3' and the reverse primer 5'-TAGTTTCAAATGAGGCGCAAG-3' and using the PCR CORE kit (QIAGEN, Mississauga, Canada). Each reaction product was amplified by PCR as follows: 1 cycle at 95°C for 5 min; 30 cycles at 94°C for 1 min, 61°C for 1 min, 72°C for 1 min; and 1 cycle at 72°C for 5 min.

2.2.4. Results

2.2.4.1. Disease latency.

Neonate mice of three different strains (BALB/c, NFS, and FVB/n) were inoculated with similar doses of the two Graffi MuLV variants GV-1.2 and GV-1.4 (see Materials and Methods). Leukemias appeared with an incidence of 99%. About 1% of the mice suffered from hindlimb paralysis and were not included in the study. Figure 2.1 shows the cumulative survival curves in the three different stains of mice inoculated with GV-1.4 (Fig. 2.1A) and GV-1.2 (Fig. 2.1B).

After the inoculation of GV-1.2 (Fig. 2.1B), 50% of the mice showed signs of disease at day 68 for the FVB/n group, day 79 for the NFS group, and day 81 for the BALB/c group. After the inoculation of GV-1.4 (Fig. 2.1A), 50% of the leukemias were detected

after 139, 158, and 181 days in FVB/n, NFS, and BALB/c mice, respectively. Thus, GV-1.2, which contains two direct repeats in the LTR, induces leukemias with a shorter latency than GV-1.4, which contains only one direct repeat as previously described (Ru, et al. 1993). Uncorrelated with the viral variant inoculated, the latency period is also significantly dependent upon the mouse strain (as measured by a log rank test; $P < 0.05$), reflecting the effect of the genetic background on the disease development. Leukemias developed more quickly in FVB/n mice than in NFS and BALB/c mice, the latter strain showing the longest latency period.

2.2.4.2. Gross pathology of leukemic mice

During the life span of the inoculated mice, no sign of pathology was detected until a few days before death. According to the Bethesda proposals for classification of neoplasms (Kogan et al. 2002, Morse et al. 2002), two groups of mice were obviously distinguishable, one presenting signs of lymphoid leukemia (thymoma and/or lymph node enlargement, no anemia) and one presenting signs of nonlymphoid leukemia (severe anemia, absence of lymph node or thymoma). The gross pathology of the leukemic mice is presented in Table 2.1. From these results, two major conclusions can be drawn. First, GV-1.2 induces more lymphoid leukemias than GV-1.4. Second, the genetics of the mice greatly influence the results: BALB/c mice are more susceptible to lymphoid leukemias, whereas FVB/n mice produce preferentially nonlymphoid leukemias marked by severe anemia. Eighteen percent of the mice showed signs of both lymphoid and nonlymphoid leukemias. Their mixed phenotypes were further elucidated by flow cytometry and by the analysis of the expression of some lineage-specific genes (see below). The results reflect the complexity of the leukemias induced by the Graffi retrovirus.

2.2.4.3. Flow cytometry.

The immunophenotype of the leukemic cells present in the different hematopoietic organs of each mouse was analyzed using flow cytometry. The leukemic population was identified according to the enlarged size of the cells and their high proportion compared to that of normal control mice.

The results shown in Table 2.1 and Fig. 2.2A reveal that the Graffi murine retrovirus is able to induce leukemias arising from all of the hematopoietic lineages: the lymphoid lineage (B and T cells), the myeloid lineage, and the erythroid and megakaryocytic lineages. Among the 108 mice analyzed, 53 leukemias of lymphoid origin were detected; among these were 35 T-cell and 18 B-cell leukemias. The other 55 leukemias had nonlymphoid characteristics and included 9 myeloid, 15 megakaryocytic, and 31 erythroid leukemias.

In BALB/c mice, GV-1.2 induced almost exclusively T-cell leukemias (91%), whereas GV-1.4 induced also B-cell (28%), myeloid (15%), and erythroid (10%) leukemias. Similar results were obtained with the NFS and FVB/n strains, although the percentage of T-cell leukemias was reduced in these strains for both variants. The broadest spectrum of leukemias is induced with the NFS strain. These results correlate well with the observations of gross pathology in mice. They confirm that the two Graffi variants induced different patterns of leukemia and that the genetics of the mice influence the results.

The immunophenotypic analyses also confirmed the existence of complex phenotypes, either mixed or biphenotypic (Kogan et al. 2002, Morse et al. 2002). Mixed leukemias (two different leukemic populations in the organism) were exclusively erythroleukemias combined with T-cell leukemias. Biphenotypic leukemias (a unique leukemic population) were essentially lymphoid, expressing specific markers of the myeloid lineage.

2.2.4.4. Global results of the induced leukemias.

To complete and confirm the gross pathology and the immunophenotypic analyses, molecular biology tools were applied to RNA and DNA isolated from each tumor (Fig. 2.3 and 2.4). For each leukemia, the flow cytometry results and the molecular characterization of the tumor RNA and DNA were compared. These global results for each type of leukemia are described below.

2.2.4.5. T-cell leukemia.

The T-cell leukemias induced by the Graffi MuLV are characterized by the presence of an enlarged thymus and spleen, often accompanied by enlarged lymph nodes, with no anemia. They were identified based on four specific markers: CD3, CD90 (Thy1.2), CD4, and CD8a. All these surface molecules are expressed by thymocytes and mature lymphocytes (Dialynas et al. 1983, Ledbetter et al. 1980.). A total of 31 T-cell leukemias were induced and analyzed. In BALB/c mice, GV-1.2 generated more than 90% of T-cell leukemias (Table 2.1), but this incidence was largely reduced in NFS (48%) and even more in FVB/n (28%) mice. For each mouse strain tested, GV-1.4 corresponded to an incidence of T-cell leukemias lower than that induced with GV-1.2 (Table 2.1).

Five groups of leukemias could be distinguished based on the expression of CD4 and CD8 markers: $CD4^- CD8^-$, $CD4^+ CD8^+$, $CD4^+ CD8^{lo}$, $CD4^+ CD8^-$, and $CD4^- CD8^+$. The two predominant forms were $CD4^+ CD8^-$ and $CD4^+ CD8^+$, representing 12 and 10 cases, respectively, (Fig. 2.2A).

Rearrangements of the TCR β gene were analyzed by Southern blotting (Fig. 2.3A and Table 2.2). Fifteen of the 31 T-cell leukemias analyzed had rearrangements of the TCR β locus. Rearrangement of the TCR β chain is initiated at the double-negative stage, and thus, we expected that the more immature phenotypes, $CD4^- CD8^-$ (double-negative stage) and $CD4^+ CD8^+$ (double-positive stage) would be those that retained a germ line configuration. Three out of four $CD4^- CD8^-$ and four out of seven $CD4^+ CD8^+$ leukemias were indeed unrearranged at the TCR β locus. But in fact, 50% of the more-mature-phenotype ($CD4^+ CD8^-$) leukemias had also a TCR β locus in germinal configuration. Thus, the TCR β locus rearrangement and the several phenotypes observed suggest that the retrovirus does not homogeneously block the differentiation pathways of the leukemic cells. No TCR β rearrangement was found in myeloid or in nonlymphoid leukemias (Fig. 2.3A and data not shown).

Recombination of the JH locus, which is specific to the B-cell lineage, was also tested, and atypically, we found that some T-cell leukemias also had rearrangements of this locus (Table 2.2). No resting B cells (less than 5%) were detected by flow cytometry in the spleen of these animals, excluding the possibility that the rearrangements observed corresponded to normal B cells (Fig. 2.2B, panels a and b). More curiously, other T-cell leukemias were biphenotypic and expressed the myeloid marker CD11b. This occurred in

five cases, two with a CD4⁺ CD8⁻ phenotype and three with a CD4⁻ CD8⁻ phenotype. Such a leukemia is illustrated on Fig. 2.2B, panel c. An aberrant abundant expression of MPO was detected in 10 tumors that were not biphenotypic (Fig. 2.4A, lane 10; Table 2.2; also data not shown), although no myeloid cells were detected by flow cytometry in the spleen of these animals. Some leukemias (Ledbetter, et al. 1980.) expressed the erythroid-specific FOG gene (Fig. 2.4B, lane 26; also data not shown), but all of them were mixed T-cell/erythroid leukemias arising predominantly in the spleen.

2.2.4.6 B-cell leukemia.

The B-cell leukemias were preferentially induced by GV-1.4 in NFS and FVB/n mice and were particularly recognizable by the presence of enormous lymph nodes and enlarged spleen but normal thymus and no anemia. They were classified as B-cell leukemias based on the presence of the surface antigen B220/CD45R, which is present from the pro-B-cell stage through the mature and activated B-cell stages, and the absence of other restricted specific lineage markers (Allman, et al. 1999, Hardy et al.. 2000). Some leukemias expressed progenitor markers at variable levels, and thus, six leukemias expressed Sca-1 and two expressed c-Kit (CD117). Rearrangements of the JH locus which occur at the pro-B-cell stage (Allman, et al. 1999) were also analyzed. Table 2.2 summarizes both the immunophenotype and the molecular characteristics of the 15 B-cell leukemias. Figure 2.3B shows the Southern blot analysis of the heavy chain immunoglobulin (JH) gene rearrangement for the B-cell leukemias in comparison with other types of leukemias and controls. Twelve out of the 15 B-cell leukemias have rearrangements of JH (Fig. 2.3B, lanes 1 to 6, 9 to 10, 12, and 14 to 16). Thus, the majority of the B220⁺ leukemias had rearrangements of JH except for those induced in NFS mice by GV-1.2 (Table 2.2).

A number of these B-cell leukemias showed a complex phenotype. First and surprisingly, 8 of the 15 B-cell leukemias expressed very high levels of MPO RNA (Fig. 2.4A), although no myeloid cells were detected by flow cytometry as was the case for some T-cell leukemias. Second, two of them were biphenotypic since they were B220⁺, had rearrangements of JH, and expressed the myeloid marker CD11b, although only one expressed MPO. Surprisingly, none of the B-cell leukemias showed TCR β rearrangement. Thus, we presume that the Graffi targeted B-cell progenitors are more committed than the

targeted T cells. JH and TCR β rearrangements may also be regulated by two separate mechanisms implicating different genes.

2.2.4.7. Myeloid leukemia.

Myeloid leukemia was recognizable by an enlarged spleen, mild anemia, slightly enlarged lymph nodes and no thymus enlargement. To immunophenotype these leukemias, we used two myeloid-specific antibodies: CD11b, which is expressed on both granulocytes and monocytes (Ho and Springer. 1982, Lagasse, and Weissman. 1996), and Gr-1, more characteristic of the granulocyte lineage (Fleming, et al. 1993, Lagasse, and Weissman. 1996). Double-positive CD11b⁺ Gr-1⁺ cells represent myeloid precursors largely found in the bone marrow of normal mice. Nine cases of myeloid leukemia were observed. Five of these were induced in NFS mice with GV-1.4, and four were induced in BALB/c mice (three with GV-1.4 and one with GV-1.2). Three phenotypes were obtained: CD11b⁺ Gr-1⁻ (five cases), CD11b⁻ Gr-1⁺ (one case), CD11b⁺ Gr-1⁺ (three cases).

We also tested the expression of MPO, which is synthesized during the promyelocytic stage of myeloid differentiation (46). Three CD11b⁺ Gr-1⁺, one CD11b⁺ Gr-1⁻, and the unique CD11b⁻ Gr-1⁺ tumor expressed very high levels of MPO (Fig. 2.4A, lanes 14 and 15; also data not shown). The other myeloid leukemias (mainly the Gr-1⁻ phenotype) did not express MPO at detectable levels. Surprisingly, MPO was not only expressed in the spleens of the animals suffering from myeloid leukemias. It was detected abundantly in many other leukemia cases, as described above for the T-cell and B-cell leukemias and shown on Fig. 2.4A. In total, 10 T-cell, 8 B-cell, 2 erythroid, and 1 megakaryocytic leukemia showed a high expression of this specific myeloid gene even though no CD11b- or Gr-1⁺ positive cells were detected in the spleens of these mice by flow cytometry. Interestingly, in contrast with the reports of Graffi (Graffi, A. 1957), we induced only one chloroleukemia (with GV-1.4 in an NFS mouse) in this study.

2.2.4.8. Erythroid and megakaryocytic leukemia.

These two types of leukemias induce severe anemia (hematocrit between 10% and 30%), an enlarged spleen, and often an infiltrated liver. The Ter119 (glycophorin A) and

CD71 (transferrin receptor) cell markers were used to immunophenotype the erythroid leukemias. Ter119 and CD71 are expressed from the proerythroblast through erythrocyte stages (Kina, et al. 2000, Lok and Loh. 1998). Megakaryocytic leukemias were identified with the use of CD41 (glycoprotein IIb), which is specific to platelets, megakaryocytes, and early hematopoietic progenitors (Uzan et al. 1991.). The megakaryocytic leukemias analyzed in this study also expressed high levels of cKit (Fig. 2.2A, panel e) and were negative for Sca-1.

These two types of nonlymphoid leukemias developed primarily in FVB/n mice (Table 2.1). In this strain, we induced 72% and 36% of Ter119⁺ leukemias with GV-1.2 and GV-1.4, respectively, and 48% of CD41⁺ leukemias with GV-1.4. Complex (mixed or biphenotypic) erythromegakaryocytic leukemias were also induced, including two with GV-1.2 and two with GV-1.4. Some mixed erythroid and T-cell leukemias also developed (as described above).

Since megakaryocytic leukemias are rarely observed in murine models of retrovirus-induced leukemias and since CD41 is also expressed at the surface of multipotent stem cells, the expression of several genes highly expressed in the megakaryocytic lineage (Pereira, R. et al. 1999, Tsang et al. 1997) was analyzed by Northern blotting (Fig. 2.4B). As expected, the results indicate that in all the erythroid and megakaryocytic leukemias, *Fog1* is expressed highly and at a much higher level than in the other types of leukemia and controls. *Fli1* also is abundantly expressed in erythroid and megakaryocytic leukemias compared to the other types of leukemias and controls (Fig. 2.4B). These two genes cannot be used to distinguish the erythroid from the megakaryocytic tumors. Interestingly, the tumors that expressed very high levels of *Fli1* also harbored a retroviral integration in the locus (not shown). *GpIIb* (CD41) was highly expressed in all the megakaryocytic leukemias and in mixed erythroid/megakaryocytic leukemias (Fig. 2.4B) in good correlation with their CD41⁺ immunophenotype. A few erythroid leukemias also expressed high levels of *GpIIb* mRNA although they were CD41⁻, possibly due to a poorer maturation of the protein in erythroid cells. *PF4* and *c-mpl* are both considered to be specific to the megakaryocytic lineage (Gurney et al. 1994). In our study, *c-Mpl* was found to be highly expressed in the majority of the megakaryocytic leukemias, whereas *PF4* was expressed in only one case (Fig. 2.4B). Surprisingly, few erythroid leukemias expressed these genes (Fig. 2.4B). The expression of these genes in both types of leukemias was often correlated with a strong expression of *GpIIb* (Fig. 2.4B). Thus, this

study enables us to conclude that the CD41/cKit leukemic cells derive from the megakaryocytic lineage. Furthermore, it shows that the erythroid and megakaryocytic leukemias induced by Graffi are very closely related.

2.2.4.9. Integrity of the U3 enhancer region.

The U3 region of the proviruses of mature tumors was analyzed to detect possible rearrangements or duplications in the enhancer region, as it was found in previous studies with other MuLVs and especially with SRS-19.6 (Granger et al. 1999).

We PCR amplified the proviral U3 enhancer region from the genomic DNA of infiltrated spleens of leukemic mice. GV-1.2 has a 60-bp duplication of the enhancer region absent in GV-1.4, and therefore the number of duplication can be deduced from the fragment's size: 351 bp for GV-1.2 and 291 bp for GV-1.4 (Fig. 2.5). We hybridized the PCR products with a U3-specific probe to confirm the specificity of the amplifications (data not shown). The results indicate that the Graffi MuLV enhancer is not stable and acquired modifications during the replication cycles. As shown on Fig. 2.5, several fragments of different size are detected for many tumors. For each tumor, the most abundant fragment generally corresponds to the input viral variant (GV-1.2 and GV-1.4), and the additional fragments correspond to altered proviruses. The smallest fragment corresponded to the intact GV-1.4 form, which is therefore the smallest enhancer region necessary for efficient replication.

In the lymphoid leukemias, some fragments were of a size larger than that expected from the input retrovirus (Fig. 2.5A and B). PCR from GV-1.4 tumors all showed the expected 291-bp fragment but also presented fragments of larger size corresponding to enhancer modification, presumably duplications and triplications. Also, PCRs for GV-1.2 tumors all showed the expected 351-bp fragment and additional larger fragments (Fig. 2.5A and B). Interestingly, a few GV-1.2 tumors contained the lower-molecular-weight 291-bp fragment typical of GV-1.4 (Fig. 2.5A, lane 3, and B, lanes 2 to 4).

The phenomenon was more rarely observed in the nonlymphoid leukemias, since they do not show duplication or multimerization of the enhancer region (Fig. 2.5C and D). All the nonlymphoid tumors induced with GV-1.4 (Fig. 2.5C, lanes 1 to 13) showed no alteration of the provirus enhancer region; lanes 2 and 8 on Fig. 2.5C show mixed

nonlymphoid and T-cell leukemias. However, GV-1.2-induced erythroid leukemias (Fig. 2.5D) often presented enhancer alteration with a majority of additional smaller fragments corresponding to the GV-1.4 enhancer (Fig. 2.5D, lanes 3 and 5 and 6 to 8). Thus, the constitution of the enhancer region can be altered during the viral replication cycles, and this surely has an impact on tumor development.

2.2.5. Discussion

2.2.5.1. Part 1

This study on the pathogenesis of the Graffi murine retrovirus indicates that it is an extremely multipotent virus, able to give rise to leukemias from every lineage of the hematopoietic system: T-cell, B-cell, myeloid, erythroid, and even megakaryocytic lineages. However, GV-1.2 and GV-1.4 induce the disease with different latencies, and more importantly, they induce different types of leukemia. GV-1.2 induces a large percentage of T-cell leukemias, although GV-1.4 induces a much wider spectrum of leukemias, including nonlymphoid leukemias.

Graffi is extremely multipotent and induces complex leukemias. The induced leukemias were found to be heterogeneous. The leukemia from each animal harbored its particularities, and several leukemias showed complex phenotypes. This heterogeneity reflects the very high polyvalence of the Graffi MuLV. In the case of T-cell leukemias, some tumors were more complex and did not show the expected phenotypes. The TCR β rearrangement was in several cases not correlated with a more mature phenotype, and more curiously, some leukemias harbored an immunoglobulin H rearrangement. This phenomenon was already observed in leukemias induced by Moloney viral recombinants and in thymomas in AKR/J mice (Granger et al. 1999, Hanecak, et al. 1988, Herr et al. 1983, Ott, et al. 1992). Possibly, the leukemic T cells derive from very early lymphoid progenitors that still retain some B-cell potential, and the deregulation induced by the viral integration may activate the machinery of the immunoglobulin rearrangement. Some leukemias showed much more complex phenotypes that could be considered aberrant. First, these are the biphenotypic lymphoid leukemias expressing the CD11b myeloid marker. Corresponding human leukemias are widely reported in the literature (for examples, see references 7, 16, 38, 44, 49, and 51), and expression of CD11b on human

leukemic B cells is associated with unfavorable prognosis (Jensen et al. 1992, Pinto et al. 1991, Tassies et al. 1995). Some studies showed that these biphenotypic cells do exist in normal mice although they are very rare (Abrahao et al. 2003, Borrello et al. 1995, Graf et al. 1999, Lu et al. 2002). Second, this concerns also the apparently aberrant expression of MPO mainly in T-cell and B-cell leukemias but also in two erythroid and one megakaryocytic case. Since RNA was extracted from whole spleen extracts, the high MPO level can be explained in two ways. First, it could be due to a large infiltration of adherent myeloid cells (histiocytes) that were not included in the flow cytometry analysis since the cell suspension was obtained by mechanical disruption without collagenase treatment. Second, the nonmyeloid leukemic cells could exhibit promiscuous gene expression of MPO as a consequence of retroviral insertional activation of this gene. We favor the latter possibility because we could not find any signs of myeloid cell infiltration in the tumors by morphological assessment of the samples (spleen imprints). This highly aberrant expression of MPO in many leukemias is probably the main feature responsible for the higher percentage of tumors classified as myeloid leukemias observed by Ru et al. (Ru et al. 1993), who used the MPO staining as an important criterion.

Thus, the high versatility of Graffi MuLV suggests that the target cells are either very early or early committed progenitors from all the lineages. Indeed, when methylcellulose colony assays were performed with bone marrow cells from infected mice, Graffi proviral DNA could be detected in every type of colony tested (CFU-G, CFU-GM, CFU-M, CFU-E, CFU-GEMM) as early as 4 days postinfection, and by 15 days postinfection, all colonies were highly positive (data not shown).

2.2.5.2. Importance of mouse genetics.

Interestingly, each strain of mice responded differently to Graffi infection. First, the latency was different among the different strains, and second, the distribution of the various leukemias was specific to each strain of mice. BALB/c mice showed the longest latency period and preferentially developed T-cell leukemias. The NFS strain was more susceptible to B-cell leukemias. Finally, the FVB/n strain showed the shortest latency and developed mainly nonlymphoid leukemias. Interestingly, this strain of mice carries an activated form of the K-ras oncogene (Shafarenko et al. 1997) which could reduce the number of hits required for leukemogenesis.

Similar observations were made in previous studies. For example, a Moloney MuLV with a mutated core motif was shown to induce 60% of erythroleukemia cases in NFS mice but almost 100% of T-cell lymphoma cases in BALB/c mice (Yuen et al. 1992). Similarly, the recombinant virus MOL4070LTR (Moloney with the 4070A LTR) induces 46% of lymphomas in BALB/c mice and 23% in FVB/n mice (Wolff et al. 2003).

2.2.5.3. Importance of the retroviral LTR.

The importance of the retroviral LTR in disease specificity is clearly established. The Graffi retroviruses are capable to infect a very large spectrum of cell types, and the major determinant is very likely due to its LTR. Indeed, we analyzed the Graffi LTR-driven transcriptional activity in several cell lines (Barat and Rassart. 1998 b). The U3 region of GV-1.2 was found to be a very strong promoter in all the hematopoietic cell lines tested (erythroid, myeloid, and T-lymphoid). GV-1.4 LTR was globally less powerful than GV-1.2 but showed a similar wide-range specificity. Thus, the Graffi U3 region could bind transcription factors present in different cell lineages and be equally activated in these cells, in contrast to Friend and Moloney viruses, which show enhanced activity only in erythroid and T-cell lines, respectively (Barat and Rassart. 1998 b).

Indeed, DNase I footprint analyses using the LTR of GV-1.4 revealed a similar pattern of protected regions in different cell lines (erythroid, myeloid, and T-cell) (Barat and Rassart. 1998 b). Some of the LTR's specific transcription factors were identified for Graffi, and we showed that GATA-1, -2 and -3 were able to activate the promoters of both GV-1.2 and GV-1.4 (Barat and Rassart. 1998 a). GATA-1 is expressed in all erythroid cell lineages, megakaryocytes, mast cells, and multipotent stem cells and has been involved in the regulation of most erythroid genes (Orkin, 1992). GATA-3 is expressed at all stages of T-cell development (Georgopoulos, 1997, Orkin, 1992). GATA-2 has a much broader expression pattern, including myeloid and erythroid cell lineages and endothelial cells, and is essential for multilineage hematopoiesis (Orkin, 1992). Thus, the importance of the GATA members in the Graffi LTR activation can explain the large spectrum of infected cells. These proteins act in cooperation with other cofactors, such as Tal1/Scl (Lecuyer, Hoang. 2004). Interestingly, a perfect consensus site for TAL1 is present proximal to the GATA binding site. It would therefore be noteworthy to identify the factors interacting with the GATA members in different cell lines.

An interesting feature that became evident in this study is that the viral variants GV-1.2 and GV-1.4 have some differences in latency and type of induced leukemia. The latency period of GV-1.2 is shorter than that of GV-1.4 (Fig. 2.1) (Ru et al. 1993), and GV-1.2 induces a higher percentage of T-cell leukemias (Table 2.1). We sequenced the two genomes, GV-1.2 and GV-1.4, which originate from the same tumor, and we found that they are very similar except for the duplication in the U3 LTR region (GenBank/EMBL/DDBJ accession numbers: GV-1.4, AB187566; GV-1.2, AB187565) (V. Voisin et al., in preparation, Voisin and Rassart, 2006). Except for the U3 region, only 16 nucleotide differences were found along their entire genomes.

The differences in latency are not likely to be influenced by viral titer since the two variants showed equivalent titers. These variations in latency are more likely explained by the presence of two direct repeats in the GV-1.2 LTR. Probably, this allows the binding of more transcription factors, making it globally more transcriptionally powerful as shown previously (Barat and E Rassart. 1998b).

GV-1.2 induces a higher percentage of T-cell leukemias. The direct repeat encompasses the well-studied region of the LVb/CORE binding sites. The CBF and MCREP bind the CORE of Moloney in T-cells and Friend in erythroid cells, respectively (Golemis et al. 1989, Li et al. 1987, Manley et al. 1993., Speck and Baltimore. 1987). Therefore, the duplicated region contains elements that are known to be more important for activation in T cells and also erythroid cells, which correspond to the major types of leukemia induced by GV-1.2. In FVB/n mice, a high incidence (72%) of erythroleukemia is induced with a very short latency compared to that in NFS and BALB/c mice. This suggests that the duplication has a strong impact on promoter activity in erythroid cells in this particular strain.

Granger et al. (Granger et al. 1999) hypothesize that the tandemization of the enhancer region has a favorable impact on the development of T-cell leukemias. Interestingly, they found that the proviruses from lymphoid tumors induced by a SRS-19.6-Moloney recombinant virus had acquired new sequence duplication in the enhancer region. Similar sequence duplications were also observed in wild-type SRS-19.6-induced tumors and were restricted to T-cell tumors. Considering the high homology between GV-1.4 and SRS-19.6 LTRs, we verified the integrity of the proviral enhancers in mature tumors. The analysis revealed that the Graffi MuLV enhancer is not stable and fluctuates between different forms during the replication cycles, showing duplication,

multimerization, and even loss of duplication (Fig. 2.5). One could hypothesize that these alterations are favorable to tumorigenesis. Granger et al. (Granger et al. 1999) proved that the tandemization of the enhancer region of SRS-19.6 was associated with T-cell leukemias but not with other leukemias. We can conclude from this study that a multimerization of the enhancer region seems to be indeed related with lymphoid leukemia. Indeed, a much higher percentage of lymphoid leukemias is generated by GV-1.2, which contains an enhancer duplication, and most of the alterations observed with this variant yield fragments of higher molecular weight. Moreover, the minimal enhancer corresponding to the GV-1.4 form seems more favorable to the induction of nonlymphoid leukemias since most leukemias of this kind arose after the inoculation of GV-1.4. In contrast with the lymphoid leukemias, nonlymphoid leukemias do not show duplication or multimerization of the enhancer (Fig. 2.5C). Furthermore, some GV-1.2-induced erythroid leukemias showed a loss of duplication (Fig. 2.5D, lanes 3, 5, and 6 to 8).

Since GV-1.2 shows a shorter latency period in addition to inducing more T-cell leukemias, one could speculate that these leukemias develop faster than the nonlymphoid ones. This was not the case, and there was no correlation between the latency period and the type of leukemia observed in each individual litter (data not shown).

2.2.5.4. Comparison with SRS-19-6 MuIV.

The sequencing of the Graffi genomes revealed an overall homology of 97% between the Graffi and SRS-19.6 MuLVs, demonstrating that they must originate from a common ancestor. However, there are enough differences to suggest that they represent distinct viruses. The nucleotide changes are dispersed along the genomes even in well-conserved regions. Two segments (the 5' end of the POL gene and the 3' end of the ENV gene) are the most divergent regions between Graffi and SRS-19.6 (Voisin et al., in preparation, Voisin and Rassart, 2006). The SRS-19.6 virus was cloned from a transmissible system developed in China and involving passaging of the viral mixture in Kunming mice, an inbred strain developed from NIH/Swiss mice (Bundy et al. 1995). This virus is able to induce a large spectrum of leukemias in NIH/Swiss mice (myeloid leukemias, B- and T-cell lymphomas, and erythroid leukemias). This distribution is very similar to our results, in particular those obtained with the GV-1.4 variant inoculated into NFS mice (Table 2.1). However, different strains of mice were used in these two studies. As the mouse genetic

background appears to be very important, the two viruses could still give slightly different phenotypes if inoculated into the same strain of mice, reflecting the differences found in their respective genomes.

2.2.5.5. A good model for megakaryocytic leukaemia.

A high percentage of nonlymphoid leukemias were induced with GV-1.4, especially in the FVB/n strain. They consist mostly of erythroid (TER119⁺ CD71⁺) and megakaryocytic leukemias (CD41⁺ cKit⁺). Some MuLVs, such as Friend and Rauscher, are known to induce erythroid leukemias, but very few studies on MuLVs have reported megakaryocytic leukemia. Thus, Graffi MuLV seems to be a good model to study megakaryocytes. Some specific genes of megakaryocytes and erythroid lineages were analyzed to confirm the true lineage of these leukemias (Fig. 2.4B). The expression of these genes did not show any erythroid versus megakaryocytic specificity or vice versa. This demonstrates the closed relationship of the two lineages that is noteworthy and widely reported in the literature. Several examples of human erythroid leukemic cell lines from patients (Rowley et al. 1992., Tani et al. 1996) or from MuLV models (Paoletti et al. 1995, Vannucchi, et al. Paoletti. 1997) are reported to bear both erythrocyte and megakaryocyte characteristics. Moreover, a common erythromegakaryocytic progenitor, bearing the surface molecules TER119 and CD41, was identified in the bone marrow (McDonald et al. 1993, Papayannopoulou et al. 1996, Vannucchi et al. 2000). This suggests that the TER119⁺ CD41⁺ leukemias observed mostly in GV-1.4-infected FVB/n mice may correspond to this progenitor and therefore represents an interesting model to study the erythromegakaryocytic lineage.

2.2.5.6. Graffi: a myeloid leukemia-inducing virus?

All together, these results show that both Graffi variants, GV-1.2 and GV-1.4, can induce a variety of leukemias in BALB/c, NFS, and FVB/n strains of mice. This exhaustive study shows that Graffi MuLV induces complex leukemias and that several immunological and molecular markers are required to classify them properly. These results are different from our previous study on the characterization of Graffi MuLV (Ru, et al. 1993), where leukemias were considered mainly myeloid. Our earlier study (Ru, et

al. 1993) was mainly based on the observations of blood smears, spleen imprints, MPO staining, and some molecular biology analysis. We described these cases as myeloid leukemias with lineage infidelities since they also contained either immunoglobulin or TCR β gene rearrangements or both. In the present study, the high levels of MPO expression in the spleen, often in lymphoid leukemias, are probably caused by aberrant gene expression.

Erkeland et al. (Erkeland et al. 2004) have described Graffi MuLV as inducing a majority of myeloid leukemias in the FVB/n strain of mice. The virus also induces a fair proportion of other tumor types. This discrepancy in results could be interpreted by a different classification approach and by the use of different antibodies. We followed the Bethesda proposal (Kogan et al. 2002, Morse et al. 2002) to classify the leukemias. In comparison with their study, we did not use the surface antigen F4/80, as we found it redundant with CD11b (Mac-1), nor did we use ERMP-58, also specific to the myeloid lineage (de Bruijn et al. 1996). We did phenotype some of the leukemias with ERMP-58 (gift of P. Leenen) (data not shown): it gave high positivity for all the myeloid leukemias, especially the CD11b⁺ Gr-1⁺ phenotype. We tested three additional antibodies: c-Kit, CD34, and CD41. This last antibody allowed the characterization of the 48% megakaryocytic leukemias found in GV-1.4-infected FVB/n mice. Finally, we cannot rule out a possible dissimilarity in the FVB/n mouse strain origin or viral isolate.

The Graffi-induced leukemias are now well characterized and can be used as a model to study hematopoiesis and leukemic progression in mice. It is now established that SRS 19-6 and Graffi are tightly related viruses and have a common ancestor. To date, as far as we know, the high incidence of megakaryocytic leukemias induced with the inoculation of GV-1.4 into the FVB/n strain of mice was never described with any other MuLVs. The results indicate also a tight relationship between the erythroid and megakaryocytic lineages, and further studies would certainly increase the knowledge about these two hematopoietic lineages.

2.2.6 Acknowledgments

We thank Denis Flipo for his help in flow cytometry. We thank Sonia Do Carmo and Philippe Legault for their help in sequencing the two Graffi genomes.

This work was supported by grant MOP-37994 from the Canadian Institutes of Health Research and by grant 3279-01 from the Natural Sciences Engineering Council of Canada.

2.3. Résultats supplémentaires

2.3.1. Schématisation plus visuelle des différentes leucémies induites par le rétrovirus Graffi

Les résultats de la pathogenèse associée au rétrovirus Graffi étant complexe, la figure 2.6 schématise de façon plus claire le spectre de leucémies induit par GV-1.2 et GV-1.4 dans les 3 souches de souris analysées. Chaque section d'anneau représente les pourcentages des différents types de leucémie obtenus pour un variant rétroviral et une souche de souris.

Ainsi on voit clairement que 1) les souris Balb/c sont plus susceptibles de développer des leucémies lymphoïdes, 2) les souris NFS sont celles qui développent les types les plus diversifiés de leucémies et 3) la souche FVB/n est plus susceptible de développer des leucémies non-lymphoïdes. Il apparaît clairement que GV-1.2 induit plus de leucémie de type T que GV-1.4. De plus, les sections d'anneaux bicolores (représentant les leucémies mixtes ou biphénotypiques) sont nombreux, appuyant la complexité du spectre leucémique du rétrovirus Graffi.

2.3.2. Clonage et séquençage du *enhancer* dans différentes tumeurs

GV-1.2 a une région *enhancer* dupliquée par rapport à GV-1.4 (figure 2.7 A). Dans l'article, il a été montré que cette région *enhancer* du virus Graffi est instable, subissant des modifications lors des cycles de réplication: multimérisation ou perte de duplication. De plus, la multimérisation tend à être associée aux leucémies lymphoïdes et la perte de duplication aux leucémies non-lymphoïdes (section.2.2.4.9).

A partir des amplifications PCR de U3 dans des tumeurs matures induites par le rétrovirus Graffi (figure 2.5), 6 fragments ont été clonés et séquencés afin de valider les résultats (figure 2.7 B et C).

Le séquençage montre que les fragments I, II et III correspondent bien à des multimérisations de la région *enhancer* (figure 2.7 B). Le fragment II est une triPLICATION du *enhancer* obtenu à partir du LTR de GV-1.2, le fragment I est une triPLICATION du *enhancer* obtenu à partir du LTR de GV-1.4, et le fragment III une duplication du *enhancer* obtenu à partir du LTR de GV-1.4. Pour le fragment II, la région tripliQUÉE (du nucléotide 91 à 150) est la même que celle de GV1.2 initialement. Pour I et III, la région multimérisÉE est un peu plus grande (de 81 à 153 pour I et de 84 à 153 pour III) mais elle comprend toujours la région entre les nucléotides 91 et 151.

Le séquençage des fragments IV, V, VI montre bien la perte au nucléotide près de la région située entre les nucléotides 91 et 151.

En conclusion, cela confirme bien que cette région précise (du nucléotide 91 à 151) est spéciale et semble conférer un rôle favorable à la réPLICATION du virus dans les cellules lymphoïdes alors qu'elle semble défavorable à la réPLICATION dans les cellules non-lymphoïdes (voir discussion article, section 2.2.5.3; des détails sur cette région seront également données au chapitre III, section 3.2.4.7). De façon étonnante, cette région, bien qu'assez conservée selon les MuLVs (figure 2.7. A) n'est pas tout à fait identique à celle qui est répétée par les autres MuLVs (chapitre III, figure 3.6). Il est difficile d'interpréter ce fait mais on peut émettre l'hypothèse que cela est dû 1) au mécanisme de multimérisation qui nécessite des séquences nucléotidiques précises ou 2) à la sélection positive de chaque région la plus favorable pour la réPLICATION. Comme les séquences nucléotidiques varient légèrement entre les MuLVs, la région la plus favorable peut être légèrement différente.

Il serait très intéressant de construire des virus chimériques avec les nouveaux LTRs multimériques obtenus et voir l'incidence sur la latence et le type de pathologie. On s'attend à ce que les leucémies soient exclusivement lymphoïdes pour les LTRs contenant une triPLICATION. La latence devrait être également réduite puisque GV-1.2 contenant une duplication a une latence plus courte que celle de GV-1.4.

Dans un but plus large, les manipulations des LTRs permettant de cibler des populations très précises sont importantes pour le développement des thérapies géniques.

2.2.3. Résultats complémentaires sur la caractérisation des leucémies

La première phase du projet de caractérisation a consisté à observer le développement de la maladie dans son ensemble parallèlement à la caractérisation par analyse de cytométrie en flux. A cause des résultats complexes obtenus antérieurement, l'approche de caractérisation devait être complète et nouvelle. Il a été notamment vérifié que les pathologies pouvaient bien être désignées comme leucémies aiguës (contrairement à un désordre prolifératif ou à des phénomènes d'inflammation). Une leucémie est caractérisée par l'envahissement de l'organisme par des cellules blastiques d'aspect immature et qui prolifèrent beaucoup. Le développement de la maladie consiste en une phase de latence pendant laquelle peu ou pas de signes cliniques sont visibles suivie d'une phase d'explosion qui conduit rapidement à la mort de l'animal (Kogan et al. 2002, Morse et al. 2002). Tous ces critères ont été observés dans le cas des pathologies induites par le rétrovirus Graffi.

Les analyses de cytométrie en flux de la taille et densité des cellules montrent que les cellules leucémiques forment une population plus grande en densité et en taille que les populations normales (figure 2.8 A). La population leucémique peut donc être facilement repérée par rapport aux populations normales et peut donc être caractérisée par cytométrie en flux. Cela permet aussi d'observer l'envahissement de la population leucémique dans les différents organes analysés. Nous avons ainsi observé que les leucémies lymphoïdes et myéloïdes ont des capacités d'envahissement plus grandes que les leucémies non-lymphoïdes érythroïdes et mégacaryoblastiques (non montré). Cela a été confirmé par l'observation des frottis sanguins qui montrent en général la présence de plus de blastes pour les leucémies lymphoïdes et myéloïdes que les leucémies érythroïdes et mégacaryoblastiques (non montré).

Le cycle cellulaire a aussi été analysé. Cela a été réalisé par cytométrie en flux après addition d'iodure de propidium (figure 2.8 B). Toutes les rates issues de souris leucémiques montrent un taux significativement supérieur de cellules cyclantes. Cependant, aucune différence significative n'a été trouvée entre les différents types de leucémies, notamment lymphoïdes et non-lymphoïdes. Ainsi, il n'est pas possible d'expliquer le plus grand envahissement de l'organisme par les leucémies lymphoïdes par une prolifération accrue. On peut alors émettre l'hypothèse que les cellules ont des propriétés migratoires différentes. Ces propriétés sont dues à la présence de molécules

d'adhésion qui vont reconnaître d'autres molécules présentes à la surface d'autres cellules ainsi que des molécules sécrétées chimio-attractantes. Ainsi les cellules non-lymphoïdes vont répondre fortement à l'attraction des cellules du micro-environnement de la rate, alors que les cellules lymphoïdes vont pénétrer plus facilement la circulation sanguine.

Des colorations pour observer la morphologie ont également été tentées (figure 2.9). Par manque d'expertise, cela n'a pas conduit à de beaux résultats. Cependant quelques observations ont pu être faites.

Cela a permis de confirmer la morphologie blastique des cellules leucémiques (tailles du noyau et des cellules plus grandes que les cellules matures). Il a été possible d'observer que les cellules lymphoïdes et érythroïdes et mégacaryoblastiques ont des propriétés de coloration différente (non montré). Les populations lymphoïdes sont plus uniformes et plus pâles que les cellules non-lymphoïdes. La population non-lymphoïde comporte des cellules grandes et pâles mais aussi petites et foncées lorsqu'on regarde une lame d'une empreinte de rate. Pour les leucémies érythroïdes, les formes caractéristiques des cellules érythroïdes en différenciation terminale, soit normoblastes et réticulocytes étaient facilement observables parmi les cellules petites et foncées sur les frottis sanguins. De petites cellules foncées étaient également présentes dans le cas des leucémies mégacaryoblastiques. Dans certains cas, les formes de réticulocytes étaient également clairement identifiables. Ceci est en corrélation avec le fait que les leucémies érythroïdes et mégacaryoblastiques induites par le rétrovirus Graffi sont très liées (section 2.2.4.8), comportant souvent des caractéristiques biphénotypiques. On peut aussi penser que certaines petites cellules pourraient correspondre à des plaquettes géantes. Effectivement des plaquettes géantes sont produites lors de défauts de la mégacaryopoïèse (Loffredo et al. 2006). La figure 2.9 montre quelques images de chaque type de leucémie. Les images a, e et f montrent des blastes lymphoïdes en division. Les images de la leucémie myéloïdes montrent les cellules caractéristiques des promyélocytes. Les images frottis sanguins des leucémies érythroïdes et mégacaryoblastiques montrent de petites cellules de la grandeur approximative d'un globule rouge.

Finalement, des essais de transplantation de cellules leucémiques dans des souris non-immunodéficientes ainsi que des essais d'établissement de lignées cellulaires ont été tentés sans succès. Des améliorations techniques- utilisation de souris *Nude* et de milieux de culture adéquats- pourraient certainement conduire à la réussite de ces

expérimentations. Ces outils sont très importants pour l'étude des leucémies et ils ont pour avantage de combler les problèmes de manque de matériel.

2.3.4 Conclusion

La nouvelle caractérisation du rétrovirus Graffi a donné des résultats complexes. Cela a permis de comprendre et d'éclaircir les résultats antérieurs de Ru *et al.* (Ru *et al.* 1993), et de faire un lien avec ceux d'Arnold Graffi. (diversification hématologique). Ainsi, les conclusions de ces 3 études ont malgré tout des points communs, avec la description de leucémies à caractéristiques multiples, A. Graffi travaillait avec un mélange rétroviral dont 2 des éléments étaient GV-1.2 et GV-1.4. On peut imaginer qu'avec un tel mélange, les types de leucémies obtenus étaient variés et on peut donc comprendre qu'il ait été incapable d'éclaircir le phénomène qu'il a appelé 'diversification hématologique' (chapitre 1, section 1.2.5.1).

Malgré cette complexité, le rétrovirus Graffi offre la possibilité d'étudier non seulement un type de leucémies mais plusieurs, dont tous ont un équivalent chez l'humain. Grâce à cette caractérisation, il est possible, avec les marqueurs utilisés en cytométrie en flux, de reconnaître et de trier de façon spécifique chacun des types de leucémies induites par le rétrovirus Graffi afin de les étudier.

2.5. Figures et Tableaux du chapitre II

Figure 2.1

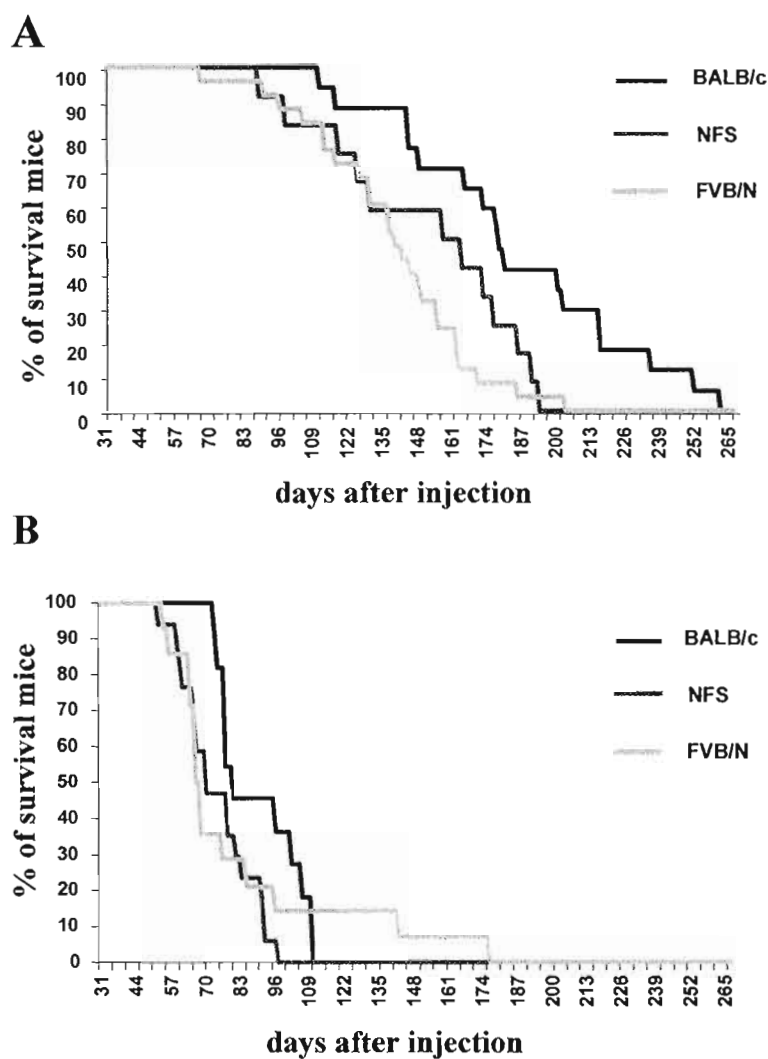


Figure 2.1: Survival analysis of newborn BALB/c, NFS, and FVB/n mice injected with the molecular clones GV-1.4 (A) and GV-1.2 (B). The animals were sacrificed when they showed signs of advanced disease. The number of mice in each group is indicated in Table 2.1.

Figure 2.2

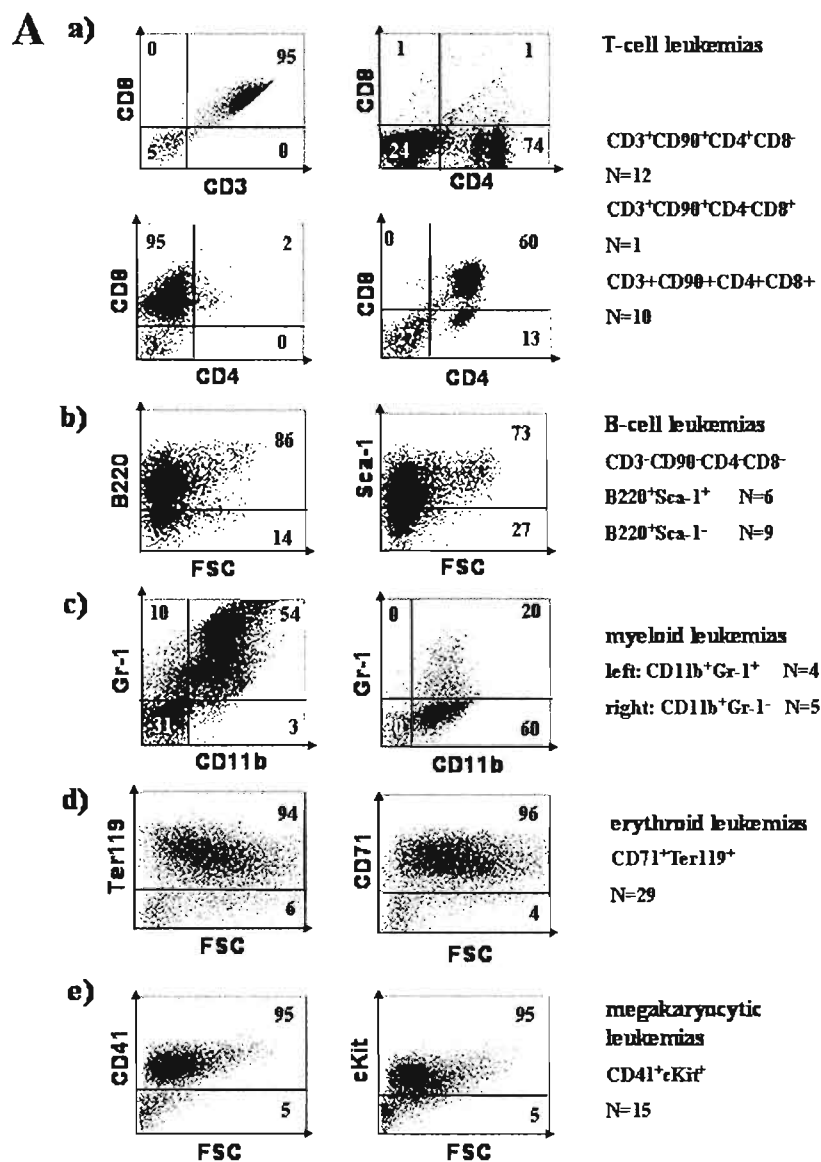


Figure 2.2: Flow cytometry analyses of several types of leukemias. (A) Flow cytometry of the most typical leukemias. Panels: a, T-cell leukemia; b, B-cell leukemia; c, myeloid leukemia; d, erythroid leukemia; e, megakaryocytic leukemia. (B) Flow cytometry of several atypical T-cell leukemias. The horizontal and/or vertical bars on each plot define the positive cells in comparison with isotype controls. N indicates the number of mice sharing the same phenotype (Table 1). The percentage of each cell population is indicated in each quadrant. The configuration of the TCR β and JH loci is indicated. Abbreviations: G, germ line; R, rearranged; FSC, forward side scatter.

Figure 2.3

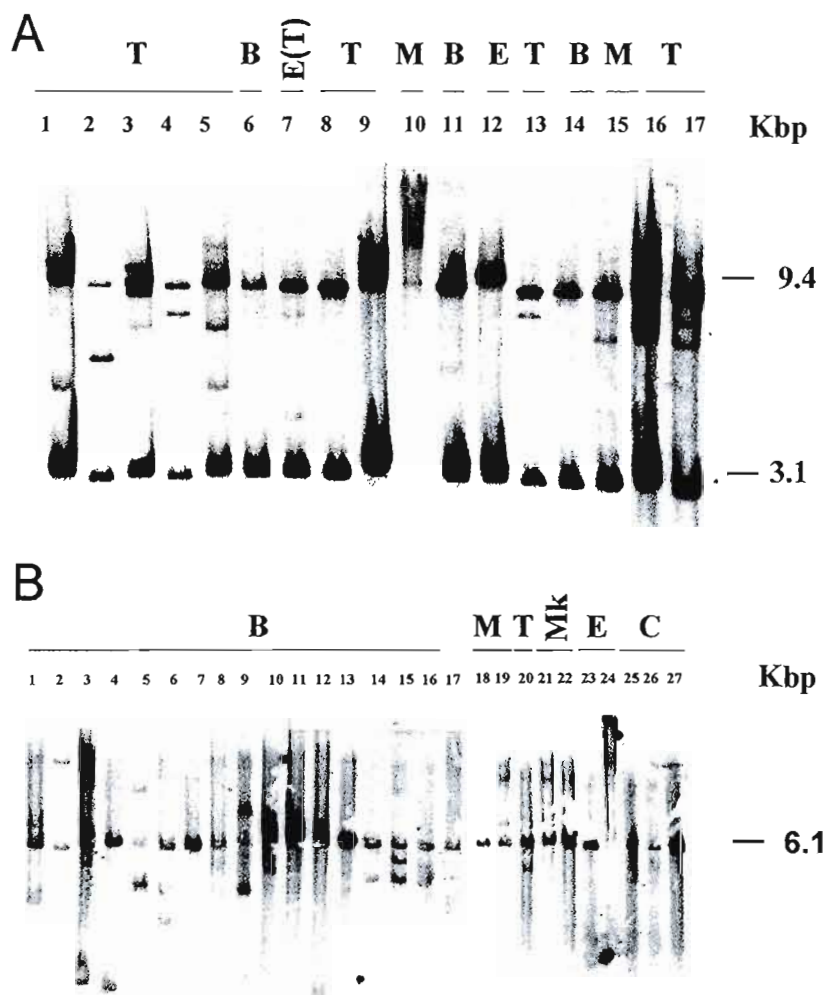


Figure 2.3: Southern blot analysis of TCR β (A) and immunoglobulin heavy chain (B) gene rearrangement. The type of leukemia is indicated above each lane. Abbreviations: B, B cell; M, myeloid; T, T cell; Mk, megakaryocytic; E, erythroid; C, controls. The molecular weights of the germ line fragments are indicated

Figure 2.4

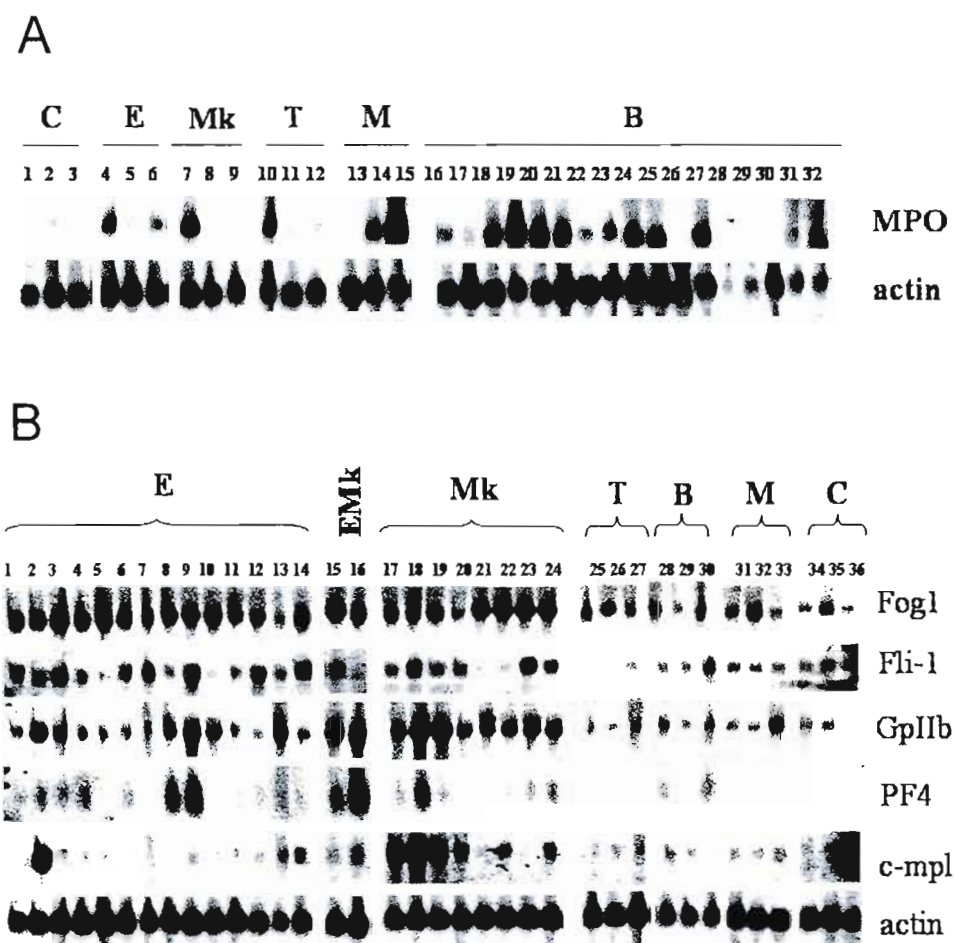


Figure 2.4: Northern blot analysis of some specific genes of the myeloid, erythroid, and megakaryocytic lineages. Total RNA was extracted from the spleens of leukemic and normal mice. (A) Analysis of the MPO gene. (B) Analysis of some genes of the erythroid and megakaryocytic lineages. Abbreviations: E, erythroid leukemias (Ter119⁺CD41⁻); EMk, erythroid-megakaryocytic leukemias (Ter119⁺CD41⁺); Mk, megakaryocytic leukemias (Ter119⁻CD41⁺); T, T-cell leukemias; B, B-cell leukemias; M, myeloid leukemias; C, control normal, BALB/c, NFS, and FVB/n mice (lanes 1 to 3 in panel A and 34 to 36 in panel B, respectively).

Figure 2.5

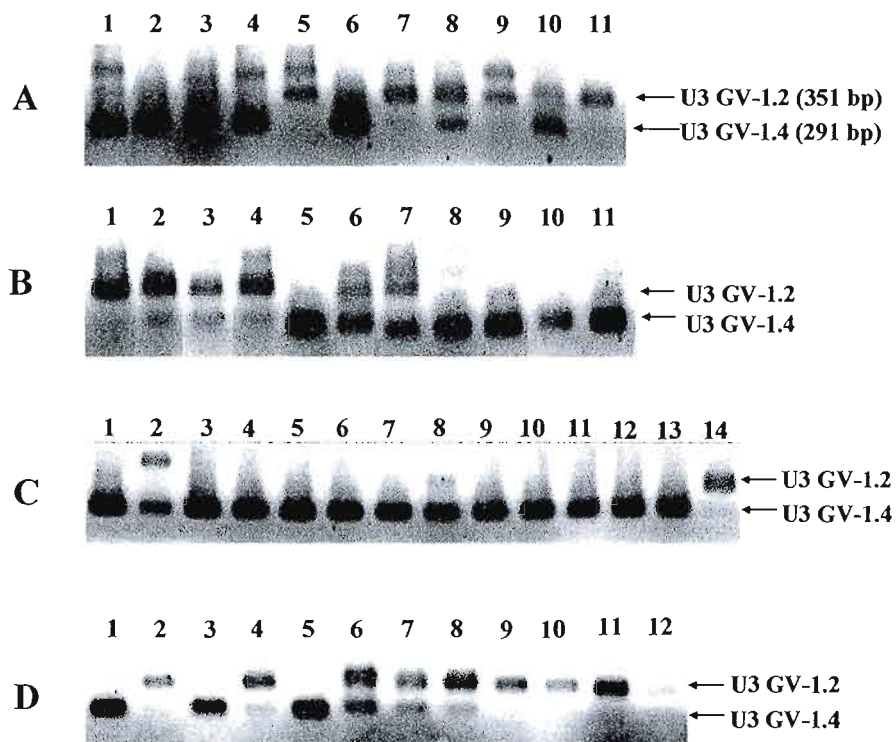


Figure 2.5: Agarose gel analysis of proviral enhancer PCR amplification in mature tumors. (A) T-cell leukemias induced with GV-1.2 (lanes 1, 2, 4, 6, 8, and 10) and GV-1.4 (lanes 3, 5, 7, 9, and 11) in BALB/c mice. (B) T-cell (lanes 1 to 6) and B-cell (lanes 7 to 11) leukemias induced with GV-1.2 (lanes 1 to 4) and GV-1.4 (lanes 5 to 11) in NFS mice. (C) Nonlymphoid leukemias induced with GV-1.4: erythroid leukemias (lanes 1 to 7) in BALB/c mice (lanes 1 to 2) and NFS mice (lanes 3 to 7); megakaryocytic leukemias in NFS mice (lanes 8 to 10); myeloid leukemias in NFS mice (lanes 11 to 13). Lane 14 is an erythroid leukemia induced with GV1.2 as a size control. (D) Nonlymphoid leukemias induced with GV-1.2: erythroid leukemias induced with GV-1.2 in NFS mice (lanes 2 to 4) and in FVB/n mice (lanes 5 to 12). Lane 1 is a leukemia induced by GV-1.4 in FVB mice as a size control.

Table 2.1

Table 1: Gross pathology and flow cytometry results of the leukemic mice

mice strain	viral variant	morphotypes				# of mice	flow cytometry				
		ane-mia	thy-moma	lymph-nodes	global		T	B	M	E	Mk
BALB/c	GV-1.2	-	10	10	L	10	10	-	-	-	-
		+	1	0	NL+L	1	-	-	1	-	-
						=11	%	91	-	9	-
NFS	GV-1.2	-	5	6	L	9	5 (3TM)	4	-	-	-
		+	0	0	NL	3	-	-	-	3	-
		+	3	1	NL+L	3	3 (TE)	-	-	-	-
		-	0	0	NL+L	2	-	1	-	1	-
						=17	%	48	29	-	23
FVB/N	GV-1.2	-	3	2	L	3	3 (1TM)	-	-	-	-
		+	0	0	NL	8	-	-	-	8 (2EMk)	-
		+	3	2	NL+L	3	1 (TE)	-	-	2 (2ET)	-
						=14	%	28	-	-	72
BALB/c	GV-1.4	-	10	8	L	11	8	3 (1BM)	-	-	-
		-	0	0	NL+L	2	-	1 (BM)	1	-	-
		+	2	2	NL+L	3	1 (TB)	1	1	-	-
		+	0	0	NL	3	-	-	1	2	-
						=19	%	64	28	15	10
NFS	GV-1.4	-	5	4	L	5	2	3	-	-	-
		+	0	0	NL	12	-	-	3	6	3
		+	2	3	NL+L	3	-	1 (E)	2	-	-
						=20	%	10	20	25	30
FVB/N	GV-1.4	-	3	3	L	3	3	-	-	-	-
		+	0	0	NL	21	-	-	-	9 (3EMk)	12
		+	0	1	NL+L	1	-	1 (BM)	-	-	-
						=25	%	12	4	-	36

L:lymphoid; NL: non lymphoid; T:T-leukemias; B:B-leukemias; M:myeloid leukemias;

E: erythroid leukemia; Mk: megakaryocytic leukemias; ():mixed or biphenotypic leukemias

Table 2.2

Table 2: global analysis of the T- and B-leukemias

Table 2. Global analysis of the T- and B-leukemias

The T-leukemias								
mice strain	viral variant	flow cytometry		molecular biology				
			#	TCRβ	Jh	#	CD11b	MPO
BALB/c	GV-1.2							
		CD4+CD8-	6	R	G	3	1+	2+
				R	R	1	-	-
				G	G	1	-	-
				R	ND	1	-	+
		CD4+CD8+	2	R	G	1	-	-
				R	ND	1	-	-
		CD4-CD8-	1	G	G	1	-	+
		CD4-CD8+	1	R	G	1	-	-
		=10						
NFS	GV-1.2							
		CD4+CD8-	3	G	G	1	+	-
				G	ND	1	-	ND
				G	R	1	-	-
		CD4+CD8+	1	G	G	1	-	-
		CD4-CD8-	1	R	R	1	+	+
			1	G	ND	1	+	+
		=6						
FVB/N	GV-1.2							
		CD4+CD8-	3	R	R	2	-	-
				G	R	1	-	-
		CD4+CD8+	1	ND	G	1	-	-
		=4						
BALB/c	GV-1.4							
		CD4+CD8+	4	R	R	1	-	-
				G	G	1	-	-
				G	R	1	-	+
				R	ND	1	-	+
		CD4+CD8	1	G	R	1	-	+
		CD4-CD8-	1	G	R	1	+	-
		CD4+CD8+	1	G	R	1	-	-
		=7						
NFS	GV-1.4							
		CD4+CD8-	1	R	R	1	-	+
		CD4+CD8+	1	G	R	1	-	+
		=2						
FVB/N	GV-1.4							
		CD4+CD8-	1	R	R	1	-	-
		CD4+CD8+	2	G	R	1	-	-
				R	ND	1	-	-
		=3						
The B-leukemias								
NFS	GV-1.2							
		B220+	4	G	R	3	-	3+
				G	G	1	-	-
		=4						
BALB/c	GV-1.4							
		B220+	3	G	R	2	-	2+
				G	R	1	+	+
		B220+ CD4+	1	G	R	1	-	+
		=4						
NFS	GV-1.4							
		B220+	6	G	G	6	-	1+
		=6						
FVB/N	GV-1.4							
		B220+	1	G	R	1	+	-
		=1						

R: rearranged locus, G: germinal configuration, ND: not determined
#: number of mice

Figure 2.6

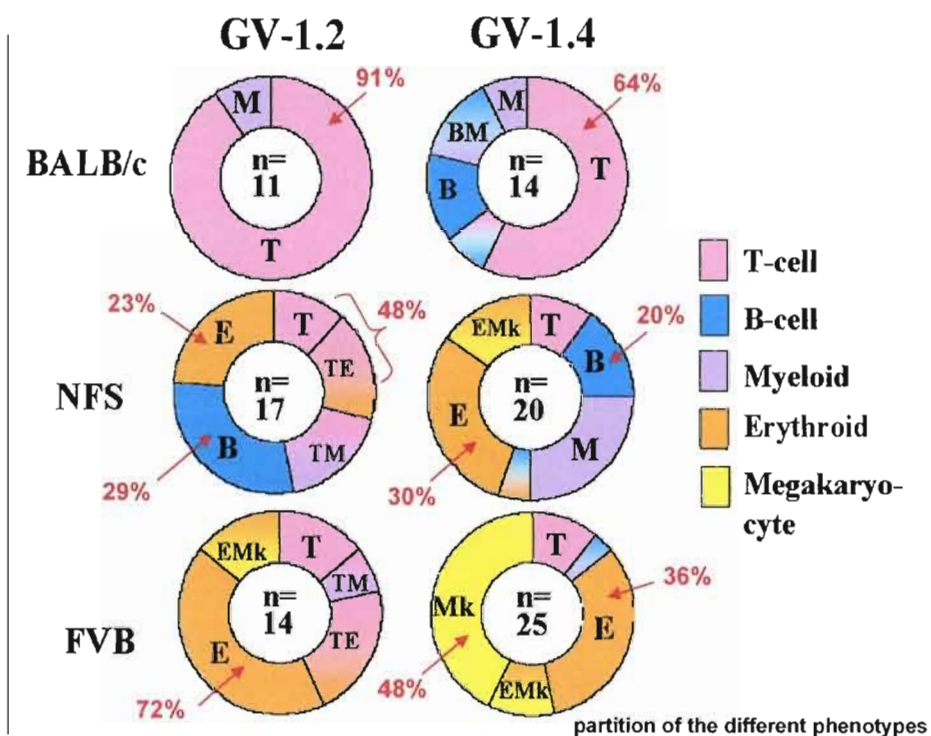


Figure 2.6: Répartition des différents types de leucémies induites par le rétrovirus Graffi. Le code couleur indique le type de leucémie et la présence de 2 couleurs dans une section d'anneau représente une leucémie mixte ou biphenotypique.

Figure 2.7

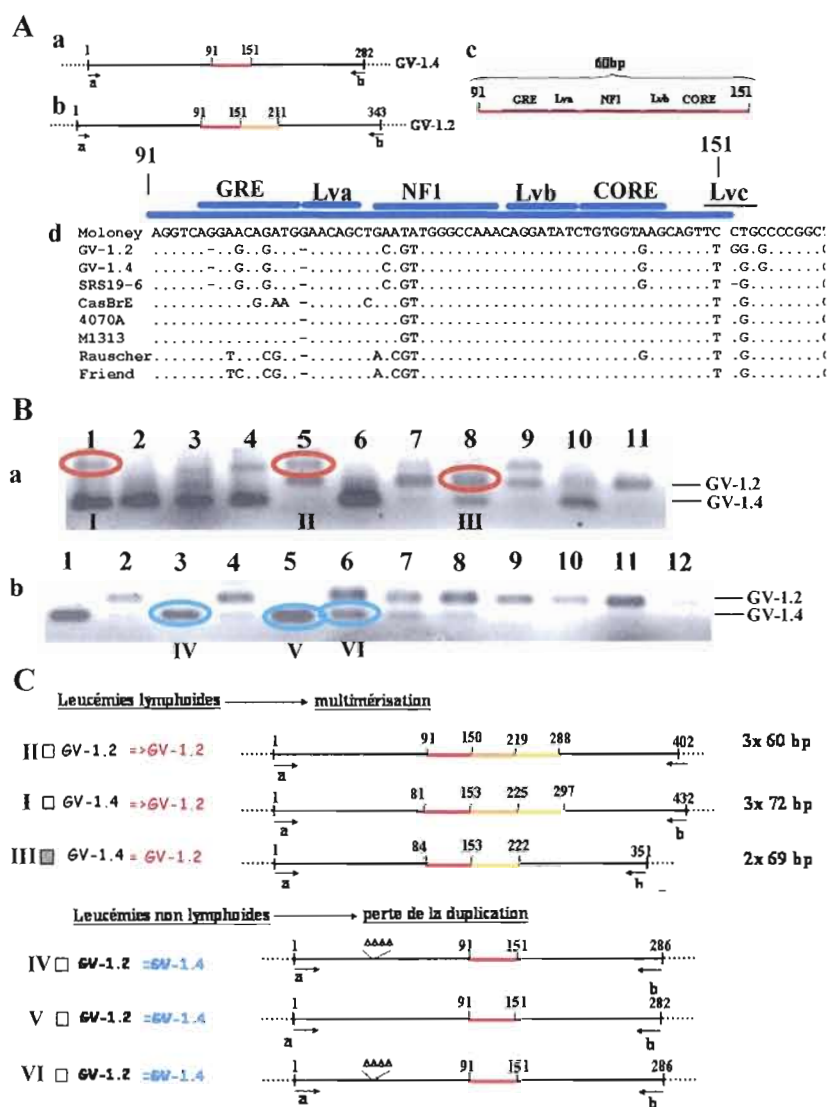


Figure 2.7: Multimérisation du enhancer, clonage et séquençage. (A) présentation de la région enhancer de GV-1.2 et GV-1.4. En rouge: région enhancer; en orange: duplication du enhancer; (a) GV-1.4; (b) GV-1.2); (c) principaux éléments du enhancer; (d) alignement de la région enhancer de plusieurs MuLVs par rapport à celui du virus Moloney. (B) Amplification PCR de la région U3. (a) répétition de la figure 2.5 A; les fragments encadrés de rouge et notés I, II, III ont été clonés et séquencés). (b) répétition de la figure 2.5 D; les fragments encadrés de bleu et notés IV, V et VI ont été clonés et séquencés. (C) Résumé des résultats du séquençage. Les fragments de U3 ont été amplifiés avec les amorces nucléotidiques notées a et b (flèches); en rouge: région enhancer; en orange: duplication du enhancer; en jaune: triplication.

Figure 2.8

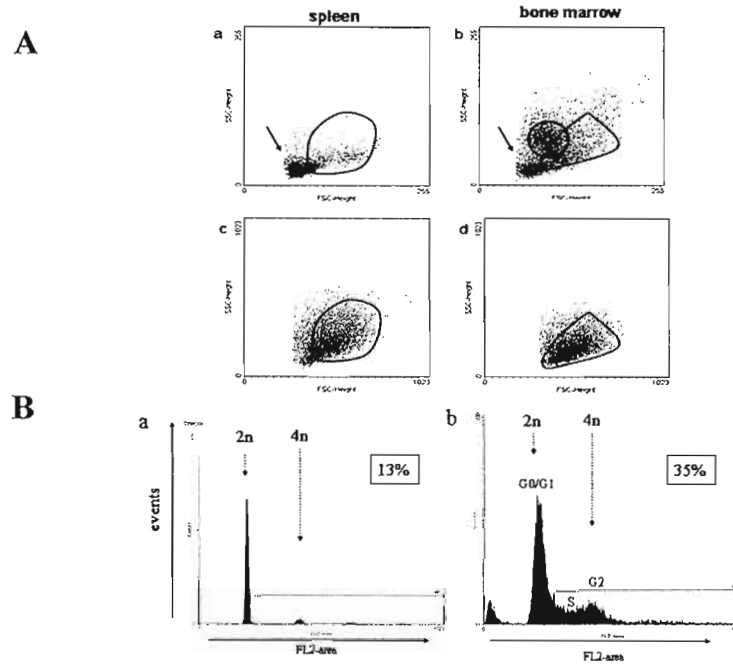


Figure 2.8: Caractérisation complémentaire des leucémies. (A) Répartition des cellules hématopoiétiques selon la taille et la densité: Forward et Side Scatters de la rate (a) et de la moelle osseuse (b) d'une souris saine et de la rate (c) et de la moelle osseuse (d) d'une souris leucémique; en rouge: la population leucémique absente pour la souris saine est encadrée de rouge; la population myéloïde normale de la moelle osseuse de la souris saine et qui a disparu pour la souris leucémique est encadré en bleu. (B) Analyse du cycle cellulaire. (a) rate d'une souris saine; (b) rate d'une souris leucémique ; Les pics 2n et 4n correspondent respectivement à la phase G1/G0 et G2; Le pourcentage de cellules cyclantes pour chaque cas est indiqué dans un encadré.

Figure 2.9

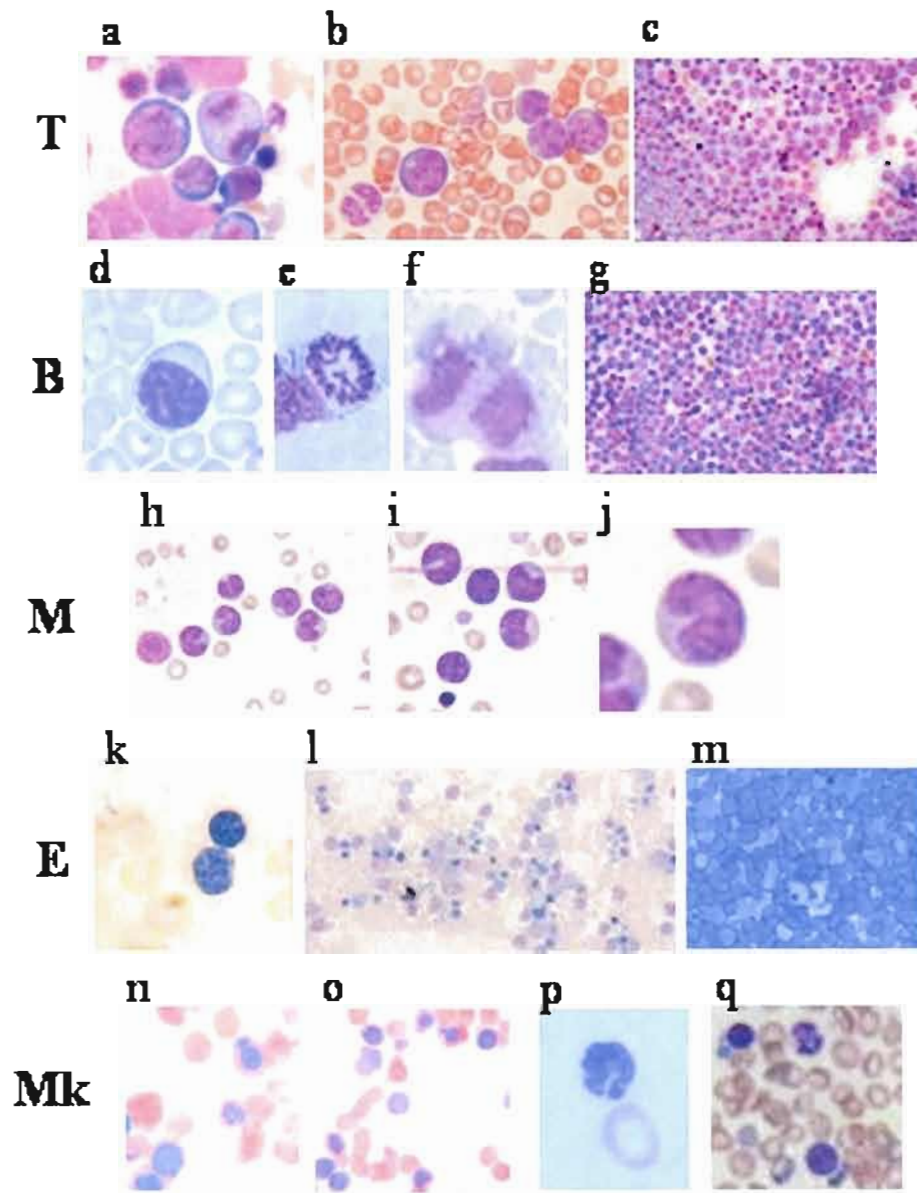


Figure 2.9:Coloration de lames (Diff-Quick). (T) leucémie T: (a)(b) frottis sanguin (blastes en division), (c) rate; (B) leucémie: (d) (e) (f) frottis sanguin, (e)(f) blaste en division, (g) ganglions); (M) leucémie myéloïde: (h) (i) (j) frottis sanguin, cellules en forme de beignet, (j) agrandissement); (E) leucémie érythroïde: (k) (l) frottis sanguin, (k) agrandissement, (m) rate; (Mk) leucémie mégakaryoblastique: (n) (o) (p) (q) frottis sanguins, (p) agrandissement.

CHAPITRE III

ANALYSES PHYLOGÉNÉTIQUES

3.1. Avant-propos

3.1.1. Introduction

Cette étude de phylogénie moléculaire compare les séquences nucléotidiques des génomes de GV-1.2 et GV-1.4 et ceux de différents MuLVs. Cette étude permet de regrouper entre eux les rétrovirus qui ont des séquences nucléotidiques qui se ressemblent le plus. On peut conclure que des rétrovirus qui forment un même groupe sur un arbre phylogénique proviennent d'un ancêtre commun et que les différences qui existent entre eux sont apparues après leur séparation et reflètent les étapes de leurs évolutions respectives.

Cette étude complète la caractérisation du rétrovirus Graffi. Cette analyse phylogénétique a été réalisée car le génome des deux variants rétroviraux, GV-1.2 et GV-1.4 (Genbank/EMBL/DDBJ numéros d'accèsion: AB187565 (GV-1.2) et AB187566 (GV-1.4)) a été séquencé entièrement. Il est intéressant de pouvoir placer ces nouvelles séquences sur un arbre phylogénique avec les autres MuLVs qui ont été séquencés. Par rapport à d'autres études de phylogénie déjà réalisées avec les MuLVs, celle-ci apporte une vision plus ample car elle inclut toutes les classes de MuLVs. Ainsi l'étude comprend des MuLVs exogènes (non présents dans le génome des souris) et endogènes (présents dans le génome de souris sauvages ou de laboratoire). Elle inclut aussi des MuLVs de classe écotrope (infectent les cellules de souris), xénotrope (infectent des cellules de mammifères autres que murines) et amphotrope (infectent des cellules de souris mais aussi d'autres espèces). L'analyse de phylogénie a fait l'objet d'un article: *Complete genome sequences of the two viral variants of the Graffi MuLV: phylogenetic relationship*

with other murine leukemia retroviruses; Véronique Voisin and Eric Rassart; Virology; 2007 May 10;361(2):335-47.

Les outils de phylogénie moléculaire sont les seuls qui permettent l'analyse de longues séquences d'ADN, ces analyses ne pouvant pas se faire manuellement. Le but de séquencer complètement GV-1.2 et GV-1.4 était premièrement de connaître les différences qui existent entre ces 2 variants. Ya-t-il des différences significatives entre ces 2 génomes autres que la duplication du *enhancer* dans le LTR de GV1.2? Cela prend toute son importance avec les résultats du chapitre II qui montrent que GV-1.2 et GV-1.4 n'induisent pas tout à fait les mêmes pathologies. Il était aussi important d'étudier les différences entre les génomes des variants du virus Graffi et celui de SRS 19-6. En effet, une étude antérieure avait noté une grande similarité (97%) entre les LTRs de GV-1.4 et de SRS 19-6 (Bundy and Fan, 1998). Comme les rétrovirus se recombinent souvent avec d'autres séquences rétrovirales lors de passages *in vivo*, la recherche de petites régions ayant recombiné étaient importantes pour retracer les étapes de l'évolution du rétrovirus Graffi. Pour ces raisons l'étude phylogénétique a dû être poussée à un niveau très détaillé.

3.1.2. Principes de base de la phylogénie moléculaire

La phylogénie moléculaire permet d'établir les relations entre différentes espèces (MuLVs dans notre cas), grâce aux similarités et dissimilarités (acquisition de mutations) entre les différentes séquences nucléotidiques (Whelan et al. 2001).

La plupart des méthodes phylogénétiques produisent des arbres sans racine: les méthodes permettent de détecter les différences entre les séquences mais ne permettent pas de savoir dans quels sens se sont produits les changements au cours du temps (pas d'information sur l'évolution). C'est le cas de l'étude phylogénétique présentée dans ce chapitre.

En effet, pour être informatif d'un point de vue évolutif, 2 méthodes existent pour enracer les arbres: 1) enracer l'arbre par désignation d'un groupe externe; 2) faire l'hypothèse de 'l'horloge moléculaire'.

- 1) Le groupe externe désigné doit correspondre à une espèce éloignée d'un point de vue évolutif de toutes les autres séquences analysées; il doit représenter théoriquement l'ancêtre commun. Aucun MuLV ne peut correspondre au critère

de groupe externe dans notre analyse. Le MuLV qui s'en rapproche le plus est HEMV (voir résultats article).

- 2) L'hypothèse de 'l'horloge moléculaire' (Zuckerkandl E. et Pauling L.) suggère que toutes les lignées (donc séquences) ont évolué à la même vitesse depuis leur divergence avec l'ancêtre commun. Cela correspond à peu de modèles et n'est certainement pas le cas dans la présente analyse. On peut effectivement penser par exemple que les rétrovirus exogènes et endogènes n'acquièrent pas les mutations à la même vitesse.

Ainsi, les arbres phylogénétiques issus de cette analyse des séquences des MuLVs sont présentés sous forme astrale et non de dendrogramme. La distance séparant deux rétrovirus sur un arbre représente le degré de similarité (ou de dissimilitude) entre ces 2 éléments. Il faut noter que certaines données d'évolution peuvent découler indirectement de cette analyse car les MuLVs sont isolés à partir de différentes souches de souris, dont l'arbre de spéciation est lui connu (voir section résultats supplémentaires).

Trois méthodes permettent de construire des arbres phylogéniques: 1) méthode des distances; 2) méthodes de parcimonie; 3) méthode du maximum de vraisemblance.

- 1) méthode des distances: elle consiste à comparer les séquences nucléotidiques 2 à 2 et le nombre de substitutions (mutations) entre les séquences est calculé. Le score entre 2 séquences représente le pourcentage de similarité. Une matrice de distance est créée avec ces scores, ce qui va permettre la construction de l'arbre phylogénique. Les branches des séquences avec les plus grands scores donc les plus proches vont être assemblées en premier et ainsi de suite jusqu'à ce que l'arbre soit complété.
- 2) méthode de parcimonie: cette méthode va générer un arbre dont la topologie est construite de façon à limiter le nombre de substitutions quand on passe d'une branche à l'autre de l'arbre. Ainsi, contrairement à la méthode de distances, plusieurs arbres sont générés par l'algorithme. Un score global est calculé en faisant la somme de toutes les substitutions nucléotidiques entre toutes les branches de l'arbre. Celui qui a le meilleur score devient l'arbre final.
- 3) méthode du maximum de vraisemblance: cet arbre est calculé de façon à ce que la topologie de l'arbre soit la plus vraisemblable compte tenu du mode d'acquisition des substitutions. Ainsi, en plus de la matrice de distances, le modèle utilise des

paramètres prédéfinis par rapport aux hypothèses d'évolution. Par exemple, dans le cas d'analyse de séquences nucléotidiques, le modèle évolutif permet de prédire une probabilité différente quand une adénosine est remplacée par une cytosine ou quand elle est remplacée par une guanine. Comme pour la méthode de parcimonie, plusieurs arbres sont générés et des scores sont calculés: l'arbre avec le meilleur score représente le plus probable et il est sélectionné pour être l'arbre final.

Les arbres générés par les différentes méthodes peuvent être soumis aux tests de *bootstrapping*. Le *bootstrap* est un algorithme qui consiste à modifier légèrement et de façon aléatoire les séquences analysées. Un nouvel arbre est généré à partir de chaque essai de *bootstrap*. Si le *bootstrapping* donne un score de 95%, cela signifie que l'arbre généré avant l'analyse a été retrouvé dans 95 des 100 essais de *bootstrap* et cela atteste de la robustesse de la topologie.

Il semble que la méthode du maximum de vraisemblance soit la plus adaptée à l'analyse des séquences nucléotidiques (conseil de Cédric Chauve, Département d'informatique, UQAM). Cependant, il est préférable de construire des arbres avec des algorithmes différents. Dans la présente analyse, les méthodes de distance, parcimonie et maximum de vraisemblance ont en général donné des résultats identiques.

3.2. Article: '*Complete genome sequences of the two viral variants of the Graffi MuLV: Phylogenetic relationship with other murine leukemia retroviruses*'

3.2.1. Abstract

A detailed phylogenetic analysis of two variants of the Graffi murine retrovirus, GV-1.2 and GV-1.4, showed that they are closely related to SRS 19-6 and Moloney MuLVs. Two stretches of sequence testify to the divergence between Graffi and SRS 19-6 MuLVs, one corresponding to a recombination event of Graffi MuLV with a xenotropic virus. Moloney MuLV was found more distant, particularly in the GAG region. Our study encompasses every class of MuLVs (ecotropic, amphotropic, xenotropic, polytropic) with some focus on exogenous ecotropic viruses and further adds to previous phylogenetic studies. Graffi, SRS 19-6, Moloney, Friend and Rauscher MuLVs form a cluster that

appears to share a common ancestor with the Casitas-amphotropic and -ecotropic MuLVs but are more distant to the Akv-type and xenotropic MuLVs. The analysis also revealed that the ENV region of HEMV, the prototype of the MuLV ancestor, was closely related to the corresponding region of Cas-Br-E.

3.2.2 Introduction

The Graffi murine retrovirus was first isolated in 1954 in Germany from an acellular extract of an Ehrlich sarcoma (Graffi, 1957). In 1993, we derived two molecular clones from the original mixture, named GV-1.2 and GV-1.4 (Ru et al., 1993). These two clones lead to infectious ecotropic NB-tropic viral particles. The Graffi MuLV is a highly potent leukemia inducing retrovirus as demonstrated in our recent study on the characterization of the pathogenesis of two viral variants, GV-1.2 and GV-1.4 (Voisin et al., 2006). This retrovirus was used as a tool to identify oncogenes at proximity of viral integrations in tumour tissues (Denicourt et al., 1999, Denicourt et al., 2003 and Erkeland et al., 2004).

The pathogenesis induced by Graffi MuLV is very complex. First described as a myeloid inducing virus, we reported recently (Voisin et al., 2006) that it is in fact very versatile and capable of inducing leukemias from each hematopoietic lineages: T-cell, B-cell, myeloid, erythroid and megakaryocytic. GV-1.2 has a shorter latency period and induces a higher percentage of T-cell leukemias compared to GV-1.4. Because of the differences in pathogenesis, it became important to determine the complete nucleotide sequence of the two retroviral genomes.

In this report, we present a detailed phylogenetic analysis of the two Graffi MuLV viral variants, GV-1.2 and GV-1.4 with several other MuLVs. This study enables to identify the degree of sequence similarities between the members of the MuLV family. It complements the recent characterization of the Graffi MuLV (Voisin et al., 2006), a retrovirus being the last exogenous MuLV discovered more than 50 years ago that deserved to be characterized.

Although mostly focused on exogenous ecotropic MuLVs, the study also encompasses every class of MuLVs (exogenous and endogenous; ecotropic, amphotropic, xenotropic and polytropic). Phylogenetic trees were constructed using the whole genomes but also using GAG, POL and ENV regions separately (nucleotidic and protein sequences). This

report completes previous phylogenetic analyses of exogenous MuLVs as it includes the sequence analysis of 16 full genome MuLVs and 21 in the envelope region.

As it is known that murine retroviruses often recombine during *in vivo* passages, one goal of this analysis was to detect such events in GV-1.2 and GV-1.4. We performed a whole genome analysis that was achieved by comparing sequential 600 bp overlapping genome segments. This report led to the identification of Graffi MuLV's closest related viruses and further underlined small regions significant of its evolution. We found that Graffi MuLV is closely related to SRS 19-6 and Moloney MuLVs with 97% and 88% identity, respectively. Two regions, one at the very beginning of the Pol gene and the other at the end of the Env gene, confirmed the divergent evolution between Graffi and SRS 19-6 MuLVs. Moloney-MuLV (Mo-MuLV) diverged most importantly from this group in the GAG region. These 3 MuLVs present significant amino-acid differences in the SU binding domain region even though the global 3D modelled structure does not seem affected. From this phylogenetic study, it appears that Graffi, SRS 19-6, Moloney together with Friend and Rauscher MuLVs are more related to the Californian wild mice viruses (Casitas-amphotropic and -ecotropic MuLVs) than to the Akv-type and xenotropic MuLVs. Our analysis also reveals that the ENV region of HEMV, an endogenous ecotropic virus found in *Mus spicilegus*, is closely related to that of Cas-Br-E MuLV.

3.2.3. Material and methods

3.2.3.1 Genome sequencing and sequences analyses

The Graffi molecular clones were derived in our laboratory (Ru et al., 1993) from a cell line chronically infected by the Graffi-MuLV (Natalie Teich, Imperial Cancer Research Fund, London, England). The DNA of two complete infectious molecular clones of GV-1.2 and GV-1.4 was sequenced at the Canadian Molecular Research services Inc. (Ottawa). The method used is based on the Sanger's dideoxy sequencing method using biotinylated oligonucleotides as primers. GV-1.2 was sequenced on both strands and GV-1.4 was sequenced on one strand. Every sequence ambiguity was resolved by sequencing the region again including the complementary strand. Each genome of GV-1.2 and GV-1.4 was covered by many overlapping contigs. The sequences have been submitted to

Genbank/EMBL/DDBJ under accession numbers AB187565 (GV-1.2) and AB187566 (GV-1.4).

3.2.3.2. MuLVs' sequences

Search for relevant MuLVs' sequences to be included in this study was done using a BLAST search of GV-1.2 and GV-1.4 sequences against databases (NCBI, <http://www.ncbi.nlm.nih.gov/blast/>), and from other published studies. A list including all the MuLVs present in this study with their respective accession numbers is presented in Table 3.1. Editing of the sequences was performed using The Manipulation Suite Server (Stothard P., <http://www.ualberta.ca/~stothard/javascript/>), DNASTar (Fairlamb, University of Dundee) and DAMBE 4.2.13 (Xia X. <http://aix1.uottawa.ca/~xxia/software/software.htm>).

3.2.3.3. Multiple alignments

The nucleotide alignment of the whole-length genomes was achieved using ClustalW (Thompson JD, Thompson et al., 1994) available at the NPS@ server (<http://www.pbil.ibcp.fr>). Alignments of nucleotide sequences from Gag, Pol and Env were conducted using ClustalW available in the DAMBE 4.2.13 package (X. Xia, see above) which ensures a conservation of the amino acid codon usage in the nucleotide alignment. Protein sequence alignments were also completed with DAMBE. The output alignments were first scrutinized to detect any misaligned segments and corrected manually if necessary. LTR sequence comparisons were performed using ClustalW available at the NPS@ server (<http://www.pbil.ibcp.fr>).

3.2.3.4. Phylogenetic analyses

Phylogenetic trees were constructed using the Parsimony method of the PHYLIP package v3.6b (J. Felsenstein, <http://www.evolution.genetics.washington.edu/>; Kuhner and Felsenstein, 1994 and Felsenstein, 1996) and PHYML 2.4.4 (Guindon S., Gascuel O., <http://www.atgc.lirmm.fr/phyml/>, Guindon et al., 2005). Default parameters were used for

both methods. Bootstrap analyses were systematically performed using 1000 trials. Only branches of the built trees carrying a bootstrap value equal to or superior to 85% (850/1000) were considered as realistic. All trees were visualized using the TreeView (Win32)1.6.6 program (R. Page <http://www.taxonomy.zoology.gla.ac.uk/rod/treeview.html>) and edited using Microsoft Powerpoint. Branches with bootstrap values lower than 50% (500/1000) were erased from the trees and branches with bootstrap values lower than 85% were indicated by dashed lines. Similar results were obtained with the parsimony (PHYLYP) and the maximum likelihood (PHYML) methods. Thus, only the results obtained with PHYML are shown.

3.2.3.5. RBD structure modelling

The ecotropic Friend MuLV receptor binding domain (1aol.pdb) was used as a reference molecule for homology modeling of the corresponding region of GV-1.4, SRS 19-6, Moloney, Akv, 1313A and DG-75 MuLVs using the program Geno3D (C. Geourjon, <http://www.pbil.ibcp.fr>, Combet et al., 2002). The 3D structures were visualised using the webware Protein Explorer (E Martz, <http://www.umass.edu/microbio/chime/pebeta/pe/protexpl/index.htm>, Martz, 2002).

3.2.4. Results

3.2.4.1. Analysis of the GV-1.2 and GV-1.4 genomes

The genome of the infectious molecular clones of GV-1.2 (8319 bp) and GV-1.4 (8259 bp) was completely sequenced (accession numbers: GV-1.4: AB187566; GV-1.2: AB187565). Many regions of both genomes were sequenced several times to avoid sequence ambiguity. Entire genomic sequence of both GV-1.2 and GV-1.4 were consequently derived from several contigs.

A global pair-wise alignment of GV-1.2 and GV-1.4 genome sequences clearly indicated that the major difference between both viruses was indeed the 60 bp direct repeat in the U3 enhancer region of GV-1.2, absent in GV-1.4. Beside this U3-specific

difference, only 16 nucleotides were found to differ along their entire genomes (9 in the LTR region, 1 in Gag, 3 in Pol, 3 in ENV), resulting only in 2 amino acid changes in POL (Voisin et al., 2006).

3.2.4.2. Global analysis

In order to position the two Graffi's variants in relation to the other members of the MuLV's family, we built a tree using MuLVs' sequences comprising whole genome length (Fig. 3.1A). Sixteen sequences were utilized (see Table 3.1), that were approximately 9 kbp in size, comprising the 5' and 3' LTRs as well as the Gag, Pol and Env genes. It is worth noting that the tree shows excellent bootstrap values on each branch, making it very realistic (100% for all but one). Its global organisation is well correlated with the viruses' tropisms. As can be seen, the tree is composed of two major groups (1 and 2 in Fig. 3.1A). The first group comprises some exogenous ecotropic retroviruses commonly used as leukemia-inducing models (group 1A and 1B) and a group of viruses isolated from wild mice in California, composed of the ecotropic Cas-Br-E (Casitas ecotropic type) and the amphotropic 1313A (Casitas amphotropic type) (group 1C). The second group includes the Akv-related (group 2A) and the xenotropic-like viruses (group 2B). Finally, HEMV, an endogenous ecotropic virus found in *M. spicilegus* occupies a central position in the tree, equally distant from group 1 and 2.

This tree (Fig. 3.1A) demonstrated that Graffi MuLV is closely related to the SRS 19-6 and Moloney MuLVs as they form one branch of the tree. The two Graffi's variants are closely linked and separated from SRS 19-6 and Moloney MuLVs that are located on different and individual branching. These positions are confirmed by maximal bootstrap values (100%). The next closest viruses to subgroup 1A are Friend and Rauscher (group 1B) and Cas-Br-E and 1313A (group 1C) MuLVs. The global alignments between GV-1.2/GV-1.4 and other genomes used to generate this tree gave identity scores of 97% for SRS 19-6, 88% for Mo-MuLV, 85% for Cas-Br-E, 83% for 1313A, and 84% for Rauscher and Friend MuLVs.

3.2.4.3. Fine scan of the genomes

The whole genomes tree (Fig. 3.1A) enabled us to establish relationships between the MuLVs. However, another interesting point when comparing virus sequences is to detect recombination events that occurred on small genome segments. Indeed, it is well known that retroviruses often recombine with endogenous viral sequences during *in vivo* passages. We thus divided the viral genome into three regions corresponding to the GAG, POL and ENV polyproteins (Figs. 3.1B, C, D). The phylogenetic analysis was conducted on both nucleotide and protein sequences. Since both gave similar results, only the nucleotide sequence analysis is presented as it contains all the phylogenetic signals. The analysis was further refined and the retroviral genome sequence was divided into smaller regions (600 bp overlapping segments, Fig. 3.2A). The resulting trees were scrutinized to detect any modifications in comparison to the whole genome tree. Fig. 3.2B provides a simplified illustration of the hyper-variable (dark) and well-conserved (white) regions between the MuLVs utilized in this study. As expected, very few differences were detected in the Pol gene, in particular in the reverse transcriptase coding region, while the highest level of divergence was observed in the SU region of the Env gene.

3.2.4.4. Analysis of the GAG region

The phylogenetic tree constructed from the sequence alignment in the GAG region (Fig. 3.1B) has a similar organisation in comparison to the whole genome tree (Fig. 3.1A). SRS 19-6 and both Graffi variants are still very closely related with only one nucleotide difference between GV-1.2 and GV-1.4 and no amino acid change. There are 7 nucleotide differences with SRS 19-6 MuLV which are dispersed along the GAG sequence, resulting in three amino acid changes.

Some MuLVs have changed their position in the tree in comparison to the whole genome tree (Fig. 3.1B, black boxes). MCF1233 and DG-75 MuLVs diverged from each other due to differences distributed throughout the entire GAG region. Mo-MuLV was distanced from the Graffi/SRS 19-6 branch (group 1A) and seemed to get closer to Cas-Br-E/1313A branch (group 1C). However, a low bootstrap value (63%) refutes the certitude of this hypothesis. Interestingly, the same organisation was also obtained in the tree issued from the protein sequence alignment but with a significant bootstrap value

(91%). The sequential phylogenetic analysis reveals that Mo-MuLV is removed from the SRS 19-6/Graffi branch on each of the three GAG segments (not shown). On the first segment (MA), it is closer to Cas-Br-E and 1313A but with a low bootstrap value (73%) (not shown). However, on the second and third segments (p12, CA, NC), it occupies a more central position within the group 1 between the Friend/Rauscher and Cas-Br-E/1313A branches (not shown).

In the entire GAG region, Mo-MuLV differs from Cas-Br-E by 157 nucleotides and 24 amino acids. Much more differences were obtained from the comparison of Mo-MuLV with Graffi and Friend MuLVs (212 nucleotides and 48 amino acids; 231 nucleotides and 48 amino acids, respectively). Therefore, the amino acid differences, less abundant between Moloney and Cas-Br-E MuLVs, confirm the excellent bootstrap value obtained for the branching of Mo-MuLV with group 1C in the GAG protein tree (not shown). We can conclude that Moloney is indeed more closely related to group 1C (Cas-Br-E/1313A) in the GAG region based on the analysis of the protein sequence and the nucleotides located in the first and second codon positions.

3.2.4.5. Analysis of the POL region

POL (about 3600 bp) is the longest retroviral polyprotein and codes for the viral enzymes. As can be seen in Fig. 3.2B, this region is well conserved, particularly in the first 2000 nucleotides, which comprise the genes coding for the protease, the reverse transcriptase, and the RNaseH. Thus, any changes in this region are significant from an evolutionary point of view. Along this region, 3 nucleotide changes are found between GV-1.2 and GV-1.4, resulting in 2 amino acid changes and there are 50 nucleotide differences between GV-1.2 and SRS 19-6 leading to 9 amino acid changes.

The POL region tree (Fig. 3.1C) is very similar to the whole genome tree (Fig. 3.1A) and to the GAG region tree (Fig. 3.1B). Mo-MuLV is located within the group 1A as it was in the whole genome tree, thus close to the Graffi/SRS 19-6 branch. MCF1233 is distant from DG-75 in group 2, as it was for the GAG region tree but different from the whole genome tree.

Seven overlapping segments were generated to perform the sequential phylogenetic analysis of the POL region (Fig. 3.2A). Since the POL region is a well conserved region

and with a tree very similar to that of the whole genome, we did not expect much variation. However, the sequential analysis of the first 3 segments of POL (the first 1500 nucleotides) reveals several interesting changes of positions within MuLVs and the relevant trees of this analysis are shown in Fig. 3.3. The typical organisation of the tree is completely lost in the first segment (Fig. 3.3A) that corresponds to the protease gene (Fig. 3.2). In group 1, the Graffi variants now stand apart from SRS 19-6 and Moloney MuLVs and several branches carry very low bootstrap values. Interestingly, 22 of the 50 nucleotide differences found between GV-1.2 and SRS.19-6 in the whole Pol region are located on the first 500 bp segment of POL (Fig. 3.3E), and 14 additional changes are located on the next 200 nucleotides of POL (not shown). Thus, 36 of the 50 nucleotide differences are located in the first 700 bp of the 5' end of Pol, and, consequently, 6 amino acid changes of a total of 7 are clustered in this region. When looking at the first segment alignment (Fig. 3.3E), Graffi MuLV seems a little closer to Mo-MuLV (17 nucleotide differences) whereas 25 nucleotide changes were found between Moloney and SRS 19-6 MuLVs. On the second sequential tree (Fig. 3.3B), SRS 19-6 is now closer to Graffi MuLV, but Mo-MuLV remains in an intermediate position within group 1. Starting from the fourth segment (Fig. 3.3D) and until the end of POL (not shown), Mo-MuLV occupies its usual position within group 1A.

HEMV, that was located between group 1 and 2 on the whole genome and the GAG region trees, joins group 1 in the first two sequential trees (Figs. 3.3A and B) with excellent bootstrap values in segment 2 (Fig. 3.3B), has a central position in the third sequential tree (Fig. 3.3C) and returns to its central position for the remaining segments (Fig. 3.3D).

Regarding group 2, a cleavage is revealed by sequential analysis as the formation of two distinct branches can be observed for the second and third segments (Fig. 3.3B and C). DG-75 is close to group 1 in the first segment of Pol (Fig. 3.3A), with only 14 nucleotide changes between DG-75 and Graffi MuLV (Fig. 3.3E). All together, the phylogenetic analysis of the 5' end of Pol, which corresponds to the best conserved region of the whole retroviral genome (Fig. 3.3.2B), shows surprising modifications (Figs. 3.3A to C), particularly for SRS 19-6 and Graffi MuLVs.

3.2.4.6. Analysis of the ENV region

The Env gene is the most variable region as it codes for the surface protein (SU) that interacts with the cellular receptors and confers the host specificity to the virus. It also codes for the transmembrane protein (TM) and the R peptide. Fig. 3.2B shows that the TM region is more conserved than the SU region. In spite of the variability of the ENV region, the GV-1.2 and GV-1.4 MuLVs differ from each other respectively by only 2 nucleotides in SU and 1 nucleotide in TM, none of which leads to an amino acid change. Globally, the Graffi ENV region is 95% identical to SRS 19-6, 85% identical to Moloney, 78% to Friend and 72% to Cas-Br-E MuLVs. In the SU region, GV-1.2 differs from SRS 19-6 by 27 nucleotides, resulting in 21 amino acid changes, with most of the differences being in the first codon's position. When compared with Moloney and Friend MuLVs, the Graffi SU differs by 75 and 106 amino acids, respectively. In the TM region, Graffi MuLV differs from SRS 19-6 by as much as 73 nucleotides but leading to only 12 amino acid differences, these being mainly in the third codon's position.

Since the retroviral envelope plays a pivotal role in pathogenesis, more sequences were compared from the databases and we included 6 additional MuLVs' ENV sequences for the analysis (marked with * in Table 3.1 and Fig. 3.1D).

As expected, the sub-divisions of the ENV tree are completely in accordance with the tropism of the viruses. Thus the ENV region tree (Fig. 3.1D) presents two major groups, corresponding to the ecotropic (group 1) and to the non-ecotropic MuLVs (group 2). Within the ecotropic group 1, Graffi MuLV is very close to SRS 19-6 and Moloney MuLVs, and related to Rauscher and Friend MuLVs (Fig. 3.1D: groups 1A, 1B). These MuLVs are distinct from the other ecotropic viruses, i.e the AKV-related MuLVs (group 1B), which agrees with previous results showing that they do not hybridize with the Eco-probe that was found specific to the Akv-related viruses (Chattopadhyay et al., 1980). The ecotropic Cas-Br-E known to bear a Casitas ecotropic (Cas-E) type ENV (Kozak and O'Neill, 1987 and Ikeda et al., 2001) is located within group 1 but the individual branching confirms its distinct origin from groups 1A and 1B. In this region, Cas-Br-E is very close to HEMV (group 1C). With respect to the non-ecotropic viruses, the amphotropic virus 1313A joined 4070A and the amphotropic-related MuLV 10A1 to form the Casitas amphotropic (Cas-A) group (group 2A). Also, MCF1233 and DG-75

MuLVs moved away from the Akv-related viruses and, with NZB9-1, MX33 and SL3-2, constitute the xenotropic-like group (group 2B in Fig. 3.1D).

The ENV region (2141 bp) was further divided into 4 overlapping segments. The first 3 segments, corresponding to the SU protein, generated very similar trees (Fig. 3.4A), also very similar to the entire ENV region tree (Fig. 3.1D). However, the fourth segment (from nucleotide 1400 to 2141), corresponding to the TM and R peptide region, was drastically different (Fig. 3.4B). There is no clear limit between the ecotropic and non-ecotropic MuLVs (Fig. 3.4B groups 1 and 2) in contrast to the ENV and SU trees (Fig. 3.1 and Fig. 3.4). MCF1233 and SL3-2 MuLVs are moving away from the xenotropic-like group (group 2B) while getting closer to the Akv-related viruses (group 1D). The group Graffi/SRS 19-6/Moloney is disassembled, and gets farther away from the Friend/Rauscher subgroup.

To better understand the modifications that occurred in this last segment, it was further fragmented. The multi-alignment analysis showed that the biggest alterations were located at the very end of TM including the R peptide. A phylogenetic tree was thus constructed with the last 150 nucleotides of the ENV region (Fig. 3.4C). Three major groups are found in this tree and interestingly, the C-terminus of the R-peptide is responsible for this clustering as shown by the amino acid comparison (Fig. 3.4D). One group ends with a consensus sequence being PxEYEP whereas the two other groups end with ECDxSRE and PExxESRE. The PExxESRE motif is more specific to the xenotropic-like viruses. Noteworthy, in this region, the two Graffi variants and RadLV carry this motif, being probably the result of a recombination event. This region of the genome of Graffi MuLV is therefore different from that of SRS 19-6 and Moloney MuLVs which carry the PxEYEP motif. Divergence in the R peptide begins after a lysine residue (** in Fig. 3.4D) located 9 amino acids after the protease cleavage site (* in Fig. 3.4D). Interestingly, protein secondary structure predicts an alpha helix and a random coiled-coil respectively upstream and downstream of this lysine residue.

Graffi, SRS 19-6 and Moloney MuLVs clustered together on the SU phylogenetic tree. However, Mo-MuLV induces different pathology compared to Graffi and SRS 19-6 MuLVs, in good correlation with the higher number of amino acid changes between Mo-MuLV and the two other viruses. We were thus interested to examine the amino acids differences in the Region Binding Domain (RBD) of SU of these 3 MuLVs to verify if any of these changes could affect the protein structure. We thus performed an alignment

of the corresponding amino acid sequence of these MuLVs (Fig. 3.5A) and we modelled the RBD structure of several MuLVs by homology with the Friend RBD structure (1aol.pdb) (Fig. 3.5B). The resolved structure of the ecotropic Friend RBD consists of a lopsided antiparallel β sandwich and a helical subdomain formed from two extended loops of the sandwich (Fass et al., 1997 and Fig. 3.5Ba). The residues in the helical domain but not those in the β sandwich form the highly variable regions (VRA, VRB, VRC) found among mammalian C-type retroviruses and responsible for the different tropisms. Thus, the helical subdomain determines the receptor specificity of the virus and its potential binding site is composed of residues that form a hydrophilic pocket at one end of the two long antiparallel helices in the RBD of Friend MuLV (Fass et al., 1997 and Fig. 3.5B, rectangle boxes).

All the ecotropic MuLVs that have been modelled (Figs. 3.5Bb–e) presents globally the same typical structural organisation and do not differ much from that of Friend MuLV (Fig. 3.5a). This is not surprising as they all bind to the mCAT receptor. However, the modelled structures of the amphotropic 1313A (Fig. 3.5Bg) and the xenotropic DG-75 (Fig. 3.5Bf) MuLVs differ and especially in the binding pocket region (rectangle box). The VRA in amphotropic and xenotropic MuLVs is considerably shorter than that of ecotropic MuLVs, whereas VRB is longer especially for the amphotropic ones. In the RBD region, GV-1.4 differs from SRS 19-6 MuLV by 12 amino acids (Fig. 3.5A). Four changes are located in VRA, 1 in VRB and 2 in VRC. When compared with Mo-MuLV, the Graffi MuLV RBD differs by 28 amino acids. Nine amino acids are located in VRA, 5 in VRB and 1 in VRC. Interestingly, the major differences between the amino-acid sequences of Graffi, SRS 19-6 and Moloney MuLVs are located in the loop1 in the VRA region (Fig. 3.5A), which is indeed a portion of the receptor binding pocket. This results only in subtle differences on the modelled structures (Figs. 3.5Ba,b,c, rectangle box) but we can suppose that changes in amino acids in this key region could affect the direct contact of SU with the host receptor.

3.2.4.7. The long terminal repeat (LTR) region

A pair-wise alignment shows that the LTR (U3/R/U5) of GV-1.4 shares 97% identity with the LTR of GV-1.2. The major difference is a 60 bp duplication in the GV-1.2 U3

region (Ru et al., 1993). Beside this difference, their LTR differs by 6 nucleotides in U3, 1 in the R region and 3 in U5.

The phylogenetic tree (Fig. 3.6A) shows that the Graffi MuLV's LTR is closely related to those of SRS 19-6 and Moloney MuLVs, to the Cas-E and Cas-A type viruses (Cas-Br-E, 1313A and 4070A), and to Friend and Rauscher MuLVs. Graffi's LTR is very dissimilar from those of the Akv-related viruses and the xenotropic-like groups. RadLV bears a LTR more similar to the xenotropic-like viruses (Fig. 3.6A). The global alignments that served to generate this tree gave identity scores of 96% between GV-1.4 and SRS 19-6 (86% with GV-1.2), 91% for Cas-Br-E and 1313A (81% for GV-1.2), 90% for 4070A (89% with GV-1.2), 88% with Mo-MuLV, Friend and Rauscher MuLVs (87%, 83%, 78% for GV-1.2).

We generated a multi-alignment of this region for both GV-1.2 and GV-1.4 with the most closely related MuLVs (Fig. 3.6B). The consensus binding sites from studies on Moloney and Friend MuLVs are represented (Golemis et al., 1990). A 2 bp mutation in the CORE region of Moloney LTR was shown to be sufficient to shift the disease outcome (erythroleukemia versus T-leukemia; Golemis et al., 1989). Thus, small nucleotide differences in this region can have a great impact on the binding of specific transcription factors. Two nucleotide differences were found between GV-1.4 and SRS 19-6 in the region encompassing the LVc binding site (Fig. 3.6B). The Graffi LTR has an imperfect CORE sequence, identical to that of SRS 19-6 and Rauscher MuLVs (Fig. 3.6B). Also, the first GRE element contains three nucleotide changes unique to Graffi and SRS 19-6 MuLVs, as well as the first NF-1 binding site which shows an A-C change in both viruses (Fig. 3.5B). The direct repeat of GV-1.2 (DR GV-1.2) is slightly different from that of Mo-MuLV (DR Mo) (Fig. 3.6C). The first GRE element is included in the DR of GV-1.2 (Fig. 3.6C) and is present twice whereas the DR Mo begins at the LVa element. The DR GV-1.2 ends within the LVc element and disrupts it although it is present twice in the DR Mo as the second GRE element.

3.2.5. Discussion

3.2.5.1. The Graffi retrovirus

We recently fully characterized the pathogenesis of the ecotropic Graffi MuLV (Voisin et al., 2006). Originally described as a myeloid-leukemia inducing virus (Graffi, 1957 and Ru et al., 1993), the Graffi MuLV induces a broad spectrum of sometimes complex leukemias (Voisin et al., 2006). Interestingly, the two Graffi variants, GV-1.2 and GV-1.4, exhibit slightly different specificities. GV-1.2 induces a higher percentage of lymphoid leukemias and with a shorter latency than GV-1.4. In a previous study, we demonstrated that the major difference between the two viruses was a 60 bp duplication in the U3 region of GV-1.2 (Ru et al., 1993). Because of the differences in pathogenesis, it became important to determine the complete nucleotide sequence of the two retroviral genomes.

The sequencing of both variants demonstrates that the 60 bp duplication in the U3 region is indeed the major difference. It is well known that the LTR and ENV regions are playing major roles in MuLV pathogenesis. Since there are no amino-acid differences in the ENV region of the two Graffi variants, we assume that the shorter latency and the higher percentage of lymphoid leukemias induced by GV-1.2 are most likely due to differences in the LTR and more specifically in the 60 bp duplication. We suppose that this duplication of the enhancer region, either because of its position and/or because of additional sequences is able to bind more activating lymphoid transcription factors (Voisin et al., 2006).

3.2.5.2. Graffi's closely related MuLVs

The phylogenetic study enables to determine that the MuLVs that are the most closely related to Graffi are the SRS 19-6 and Moloney MuLVs (97% and 88% identity respectively). The Graffi murine retrovirus was first isolated in 1954 in Germany from an acellular extract of Ehrlich sarcomas, named sarcomas 1 and 2 (Graffi, 1957), during the same period at which Moloney, Friend and Raucher retroviruses were discovered. Moloney MuLV was extracted from the so-called Sarcoma 37 (Moloney, 1960). From the results of this study, we hypothesize that the tumors that carried Graffi and Moloney

MuLVs may have been derived from a common Ehrlich sarcoma that was transplanted several times and passed between different laboratories.

The SRS 19-6 MuLV was discovered from a system of transmissible leukemia that was also developed about 40 years ago in the People's Republic of China. This work was carried in Kunming mice, a strain derived from Western laboratory mice (Bundy et al., 1995). A previous study already reported the high similarity between SRS 19-6 and the Graffi LTRs (Bundy and Fan, 1998). Albeit an apparent different history, there is no doubt that Graffi and SRS 19-6 MuLVs share a common ancestor. The differences in their genomes were acquired during multiple *in vivo* passages in different strains of mice before their molecular cloning. Two stretches of sequence, one at the very beginning of POL and the other at the end of ENV, testify of their divergent evolution from the common ancestor. Both viruses generate the same kind of pathology: SRS 19-6 induces a broad spectrum of leukemias when injected into NIH/Swiss newborn mice (Bundy et al., 1995), a pathology very similar to that of GV-1.4 in NFS mice (Voisin et al., 2006). We showed that the mouse genetic background affects the disease phenotype. Thus, we inoculated the SRS 19-6 MuLV (kindly given by Dr. Hung Fan) in BALB/c mice and the Graffi GV-1.4 virus was inoculated in NIH/Swiss mice. Although the experiment is still ongoing, the results show that the two viruses induce comparable phenotypes: lymphoid leukemia in BALB/c and a wider spectrum in NIH/Swiss mice (H. Fan and E. Rassart, unpublished results). There are some differences in amino acids between the SU of SRS 19-6 and that of Graffi/MuLVs, especially concerning key amino-acids of the loop1 located in the receptor binding pocket region. Although they do not seem to influence the global 3D structure of the RBD, it is not possible to exclude the potential importance of these differences in the pathogenesis of the two viruses (Fig. 3.5).

The sequence differences, more abundant between Mo-MuLV and the two other members of the group (Graffi/SRS 19-6) are reflected in the pathogenesis: Mo-MuLV induces exclusively T-lymphoid leukemias in NFS and NIH/Swiss mice. First of all, the LTR sequences are certainly involved in the differences in pathogenesis between Mo-MuLV and Graffi/SRS 19-6. Mo-MuLV contains a direct repeat in its LTR as does GV-1.2 (Fig. 3.6). Indeed, GV-1.2 induces a higher percentage of T-lymphoid leukemias than GV-1.4 and SRS 19-6. However, in NFS mice, GV-1.2 does not induce exclusively this type of leukemia, but a much broader spectrum of diseases (Voisin et al., 2006). The duplication in the LTR is not exactly the same between GV-1.2 and Mo-MuLV (Fig. 3.6)

and it is likely that the Mo-MuLV LTR can bind more T-cell specific transcription factors than that of GV-1.2. GV-1.4 and SRS 19-6 LTR are very similar and are probably more suitable for the binding of non-lymphoid transcription factors as it was partially shown for the GV-1.4's LTR in a previous study (Barat and Rassart, 1998).

However, Mo-MuLV recombinants carrying the SRS 19-6 LTR still induce T-lymphomas, showing that the SRS 19-6 LTR is not sufficient to confer its broad disease specificity (Granger et al., 1999). Thus, the ENV region and, in particular the SU protein contributes to explain the differences in disease phenotype. There are indeed many amino acid differences between the SU of Mo-MuLV and those of Graffi/SRS 19-6 MuLVs, especially concerning key amino acids located in VRA, VRB and VRC (Fig. 3.5). Interestingly, new insights into the determinants of disease specificity of Moloney and SRS 19-6 MuLVs are brought by a recent study of Jahid et al. (2006). First, they demonstrated that the ENV of SRS 19-6 is more powerful in inducing the production of MCF recombinants. They demonstrated that GAG and POL regions also seem to carry some determinants for disease specificity. There are definitively significant differences between the GAG and POL sequences of Moloney, Graffi and SRS 19-6 MuLVs. The results of the phylogeny study show that Mo-MuLV is more distantly related to Graffi and SRS-19-6 MuLVs on the entire GAG region with many nucleotide and amino acid changes whereas Graffi and SRS 19-6 MuLVs remain very close. The capsid protein in the GAG region is the determinant of the Fv-1 NB tropism and Graffi and SRS 19-6 MuLVs are similar in this region but they are different from Moloney-MuLV and also from other ecotropic MuLVs studied (Stevens et al., 2004). An E at position 110 of the Graffi and SRS 19-6 MuLVs' capsid protein is normally associated with B tropism (Kozak and Chakraborti, 1996). Since these viruses are both NB tropic, the Fv-1 restriction must be modulated by neighbouring residues (Stevens et al., 2004 and Lassaux et al., 2005). Candidate residues involved in this restriction might be E82, N94, N103, H106 or some combinations of these (Stevens et al., 2004). Our phylogeny study demonstrates also an unexpected divergence of the 3 viruses at the beginning of POL, in the protease coding region (Fig. 3.3), a region usually highly conserved (Fig. 3.2B). Thus, if some disease specificity determinants are located in POL, this region is probably a good candidate to be tested in chimeras experiments.

The next closest related MuLVs to group 1A are Friend/Rauscher (group 1B) followed by Cas-Br-E/1313A (group 1C). The tight related history of Friend and Rauscher MuLVs

was already reported in a previous phylogenetic study on Rauscher and other MuLVs (Khimani et al., 1997). There is a 97.6% overall sequence identity between Friend and Rauscher MuLVs and they both induce erythroleukemias. The fact that Friend MuLV is also derived from an Ehrlich tumor (Table 3.1), probably explains why this virus is close to group 1A (Graffi/SRS19.6/Mo). It is not surprising to find Cas-Br-E and 1313A so closely related even if they have different tropisms, as they were both isolated from wild mice trapped in California (Table 3.1, Gardner, 1978, Kozak and O'Neill, 1987 and Ikeda et al., 2001). The LTR sequences of these Cas-E and Cas-A MuLVs are highly similar to those of Graffi and SRS 19-6 MuLVs. Interestingly, all these MuLVs also share the ability to induce a variety of different hematopoietic neoplasms (Holmes et al., 1986, Ott et al., 1990, Bergeron et al., 1992, Bundy et al., 1995, Wolff et al., 2003 and Voisin et al., 2006). These Cas-type LTRs are also closely related to Moloney, Friend and Rauscher MuLVs (Fig. 3.6A).

This study indicates that the Casitas type viruses are more similar to the group formed by Graffi/SRS 19-6/Moloney and Friend/Rauscher than to that of Akv-related and xenotropic viruses. It is very likely that these Casitas type viruses share a remote ancestor with Graffi/SRS 19-6/Moloney/Friend/Rauscher MuLVs.

3.2.5.3. MuLVs more distantly related to Graffi

Akv MuLV is a prototype as it represents the first MuLV isolated by Gross from AK and C58 strains of mice in 1951 together with the Gross passage A isolate (Gross, 1951 and Gross, 1957). The present study shows that Akv, SL3-3 and PSR3 remain clustered in all the trees (Figs. 3.1B–D). The analysis of the sub-regions did not reveal any reshuffling in the position of these three MuLVs, confirming their common history and possibly more recent divergence as previously reported (Pedersen et al., 1981 and Nouvel et al., 1993). One report identified a nucleotide stretch in the integrase coding region of SL3-3 differing from Akv but identical to 10A1 (Lund and Pedersen, 1999). PSR3 is almost identical to Akv (Nouvel et al., 1993) except for the R peptide which allows differentiating Akv/SL3-3 from PSR3 (Fig. 3.1D).

RadLV and BM5eco are also located closely to the Akv-related viruses throughout the analysis. Interestingly, RadLV bears xenotropic-like LTR and R peptide (Figs. 3.6A and

4E) (Merregaert et al., 1987 and Poliquin et al., 1992). BM5eco which causes murine AIDS in susceptible mice was already reported to be related to the Akv-type viruses (Gaur and Green, 2003). The present study reveals that no correlation can be established between Graffi-MuLV's history and that of the Akv-related viruses group.

DG-75 is the only xenotropic MuLV included in the whole genome sequence. It remains closely linked to the Akv-related viruses on the corresponding trees (Figs. 3.1A, B, C). Interestingly, even if these MuLVs are not closely related to Graffi MuLV, they share some similarities in the R peptide region, suggesting a recombination event with a xenotropic-MuLV in the Graffi-MuLV's history after the divergence from SRS 19-6.

Noteworthy, we found three groups of MuLVs with distinct C-terminal sequence motifs in the R peptide: PExxESRE, ECDxSRE and PxEYEP (Fig. 3.4D). The cleavage of the R peptide is fundamental for the fusion activity of the envelope (Thomas et al., 1997, Rein et al., 1998, Melikyan et al., 2000 and Aguilar et al., 2003) and the cleavage site is well conserved among the MuLVs although no other function was clearly assigned to the R peptide. We observed that the well-conserved N-terminal part of the R peptide bears an alpha helix structure although the hypervariable C-terminal part has a less well defined structure (random coiled-coils). Thus, the absence of a well-defined structure correlates with the high variability of the C-terminal end of the R peptide and this could be correlated to poor functional importance contributing to the numerous phylogenetic signals.

HEMV is also distantly related to Graffi MuLV. Its distinctive central position in all the trees of this study corroborates previous studies suggesting that this MuLV could be related to the common MuLV ancestor (Tomonaga and Coffin, 1999 and Tipper et al., 2005). Noteworthy and surprisingly, Cas-Br-E moves very close to HEMV in the ENV region (Fig. 3.1 and Fig. 3.4).

HEMV is exclusively found in *M. spicilegus* (Tipper et al., 2005). In an earlier study, it was found that the Cas-Br-E ENV region was close to that of Ho-MuLV (Khimani et al., 1997), a related but different virus from HEMV (Voytek and Kozak, 1989). Similar to HEMV, Ho-MuLV was also isolated from *M. spicilegus* (Voytek and Kozak, 1989). To further evaluate the phylogenetic relationship between Cas-Br-E and HEMV ENV regions, we compared Cas-Br-E, HEMV, Ho-MuLV (M26527) and Fv-4 (Cas-E type ENV, NM01085) (not shown). Cas-Br-E is more related to Fv-4, and HEMV is more

related to Ho-MuLV. We can thus conclude that HEMV and Cas-Br-E share a common ancestor but it is ancient and may have been present before the separation of *M. spicilegus* and *Mus musculus* germ lines.

The major goal of the present analysis was to identify those viruses closely related to Graffi MuLV. We performed the analysis on MuLVs from different classes to obtain additional information. Several trees from this phylogenetic study highlight that viruses having similar sequences were often isolated from a common source. This suggests that most if not all MuLV members derive from a few numbers of ancestors. These ancestor viruses recombined and acquired mutations during in vivo and in vitro passages, either in the wild or in laboratory mice, thereby contributing to the diversity of the MuLV family. The sequential analysis is a quick method to reveal divergence between various MuLVs on small sequence segments. For example, it revealed an unexpected divergence between Graffi, SRS 19-6 and Moloney at the very beginning of POL, a region usually highly conserved. Altogether, this study demonstrates that Graffi MuLV genome sequence is similar to the other ecotropic viruses and that a single recombination event between Graffi MuLV and an endogenous virus occurred at the end of ENV. Most likely, a portion of the LTR was also involved in this recombination. Indeed, the first 36 nucleotides of the Graffi GV-1.4 LTR are identical to DG-75 LTR. It is tempting to speculate that the Graffi GV-1.2 variant, which contains a 60 bp duplication in its LTR, is most probably derived from the GV-1.4. Indeed, we and others already showed that such duplication and/or multimerisation in the U3 LTR region is often observed in lymphoid tumors derived from MuLVs infected mice (Rassart et al., 1986; Granger et al., 1999 and Voisin et al., 2006).

3.2.6. Acknowledgments

We thank Drs. Cédric Chauve, Laurent Poliquin and Benoit Barbeau for helpful discussion. We thank Sonia Do Carmo and Philippe Legault for their help in sequencing the two Graffi genomes. This work was supported by grant MOP-37994 from the Canadian Institutes of Health Research and by grant 3279-01 from the Natural Sciences Engineering Council of Canada.

3.3. Résultats supplémentaires: Distribution géographique des MuLVs

La figure 3.7 est une façon alternative de montrer les résultats de l'étude de phylogénie. Elle illustre la distribution géographique des principales souches de souris qui existent dans le monde et montre les différents MuLVs, inclus dans l'analyse de phylogénie et qui ont été isolés à partir de ces souris. Les informations nécessaires à la réalisation de cette figure ont été prises en majorité de l'article de Kozak et O'Neill (Kozak and O'Neill, 1987) qui ont étudié en détail la distribution de séquences provirales dans les souris sauvages, couvrant le continent Nord-américain, l'Europe de l'Est et de l'Ouest et le continent asiatique. La figure contient également les données apportées par l'étude de phylogénie. D'autres articles ont été des ressources précieuses: Ikeda et Kato (Ikeda et Kato 2001) Gardner (Gardner 1978), Chattopadhyay and Lander (1980) Un code couleur indique quels MuLVs ont le plus de similarité entre eux d'après l'étude de phylogénie.

En Amérique du Nord et en Europe de l'Ouest, la souche de souris la plus répandue est *Mus. musculus Domesticus* qui contient dans son génome des traces ancestrales d'infections rétrovirales sous la forme de virus défectifs (MCF). Les souris de laboratoire ont comme mère ancestrale une souris *M.m Domesticus*. Mais afin de créer des souches de souris variées et caractéristiques, les chercheurs les ont croisées avec bien d'autres espèces. Le groupe de rétrovirus Graffi-, Moloney-, Friend-, et Rauscher dérivent de ces souris de laboratoire et d'après les résultats de la phylogénie et ils dériveraient probablement d'une même tumeur d'Ehrlich (section 1.2.5.2). Les tumeurs d'Ehrlich représentent le premier modèle d'étude du cancer: elles furent développées par le médecin allemand Paul Ehrlich à la fin du 19^{ème} siècle qui a pu pour la première fois transplanter des tumeurs spontanées murines d'une souris à l'autre; les tumeurs transplantées se forment soit sous forme liquide (ascites) ou solide (sarcomes) (Cailleau et Costa , 1961).

Le rétrovirus SRS 19-6, très proche d'un point de vue phylogénique de ce groupe et notamment du virus Graffi, a été isolé en Chine à partir du complexe TSZ isolé à partir de souris Kunming dérivées de souris occidentales (section 3.2.5.2). Ainsi cela n'a pas de corrélation avec les autres souris trouvées à l'état sauvage sur le continent asiatique, soit *M.m. musculus*, *M.castaneus*, *M spicelegus* ou *M. Molossinus* (mélange de *M. m. musculus* et *M. castaneus*).

Les résultats de la phylogénie montrent que le groupe Graffi-, SRS 19-6-, Moloney-, Friend-, Rauscher- MuLVs est plus proche des virus de type amphotropique que des types liés à Akv (section 3.2.5.2). Les virus amphotropiques ont été trouvés chez des souris sauvages en Californie, Ces souris (du Lac Casitas) seraient un mélange de *M.m domesticus* et de *M.castaneus*. Ainsi on peut penser que la similarité entre les 2 groupes provient des infections ancestrales de *M. Domesticus*. Le bagage génétique de ces souris du lac Casitas comporte également des éléments de *M. castaneus*: ceci explique que Cas-Br-E porte une enveloppe écotropique de type Fv4 (CasE), dérivant probablement des infections ancestrales de *M. Castaneus* par des virus de type CasE (Chine,Thaïlande).

Les virus de type Akv n'ont aucune relation avec le groupe Graffi-, SRS 19-6-, Moloney-, Friend-, Rauscher- MuLVs (section 3.2.5.3). Ils proviennent certainement de souris de laboratoires très différentes.

Finalement, la distribution géographique montre que les virus xénotropiques se trouvent plus souvent dans des souris d'Europe de l'Est et d'Asie.

En conclusion, cela montre que la phylogénie a regroupé les virus qui ont un ancêtre commun et cela est en corrélation avec la distribution géographique des souris à partir desquelles ils ont été isolés et qui ont subi des infections ancestrales.

3.4. Figures et tableaux du chapitre III

Figure 3.1

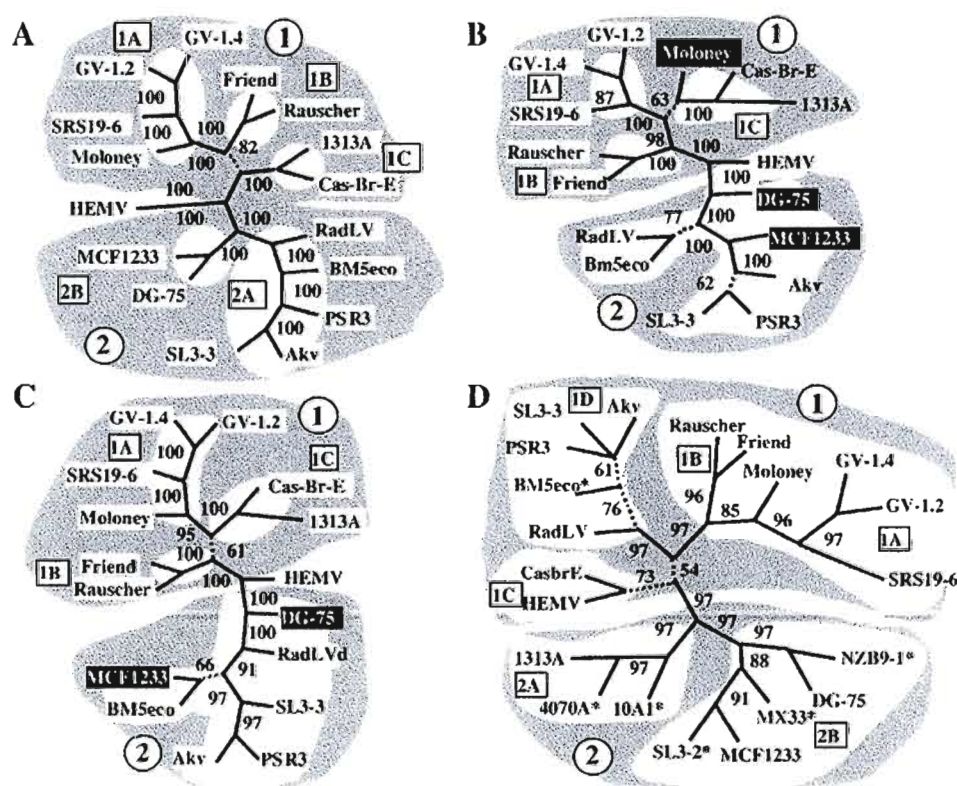


Figure 3.1: Phylogenetic analysis of several MuLV genomes. The phylogenetic trees were constructed from (A) the complete genome sequence, (B) the Gag gene, (C) the Pol gene and (D) the Env gene. Bootstrap values are expressed in percentage and are indicated for each branch of the trees. When the bootstrap value was lower than 85%, the branch was drawn with a dashed line. The MuLVs represented in black boxes indicate those that changed position in the tree in comparison with the whole genome tree (A). For each tree, the two major groups are indicated in dashed grey and numerated (1 and 2). The subgroups are also numerated (1A-D to 2A-B).

Figure 3.2

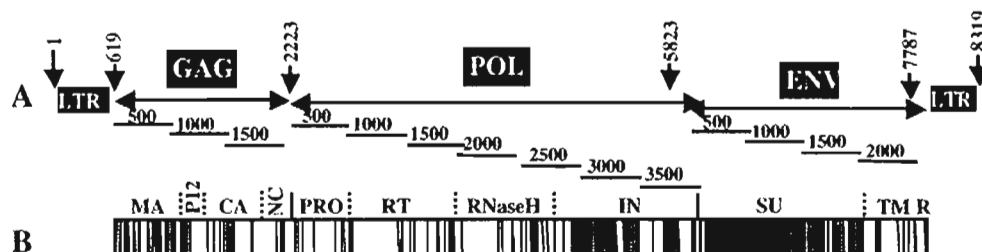


Figure 3.2: Phylogenetic analysis strategy and detection of hypervariable regions.(A) Organisation of GV-1.2 retroviral genome. Each gene and corresponding proteins are indicated. The numbers above the viral genome show the nucleotide corresponding to the start of each gene. The segments used in the sequential phylogenetic analysis are shown under the viral genome. Each segment was 600bp long with a 100bp overlap with the next one. (B) Detection of hypervariable regions in the MuLVs' genomes. The results of the multiple alignments are represented. The non-conserved regions are indicated by vertical lines.

Figure 3.3

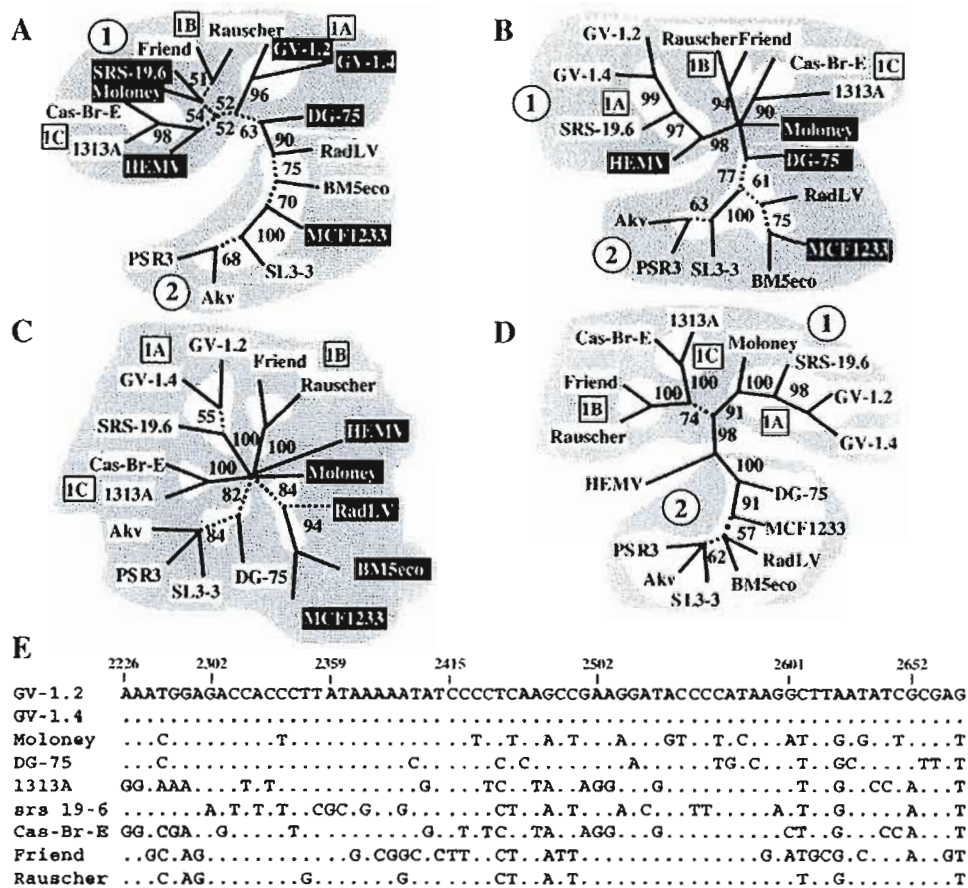


Figure 3.3: Sequential phylogenetic analysis of the Pol gene. Phylogenetic trees of the first four segments (see figure 3.1A) of the Pol gene are illustrated (A to D). Bootstrap values are indicated for each branch of the trees. When the bootstrap value is less than 85%, the branch is drawn with a dashed line. The MuLVs represented in black boxes indicate those that changed position in the tree in comparison with the whole genome tree (figure 3.2A). For each tree, the two major groups are indicated in dashed grey and numerated (1 and 2). The subgroups are also numerated (1A-D to 2A-B). (E) Multi-alignment of the first 500bp segment. Only the non-conserved nucleotides are shown and thus do not represent a continuous sequence. Some nucleotide positions corresponding to the GV-1.2 sequence are indicated to facilitate the understanding. The alignment was performed with the 9 MuLVs of group 1 from the whole genome tree.

Figure 3.4

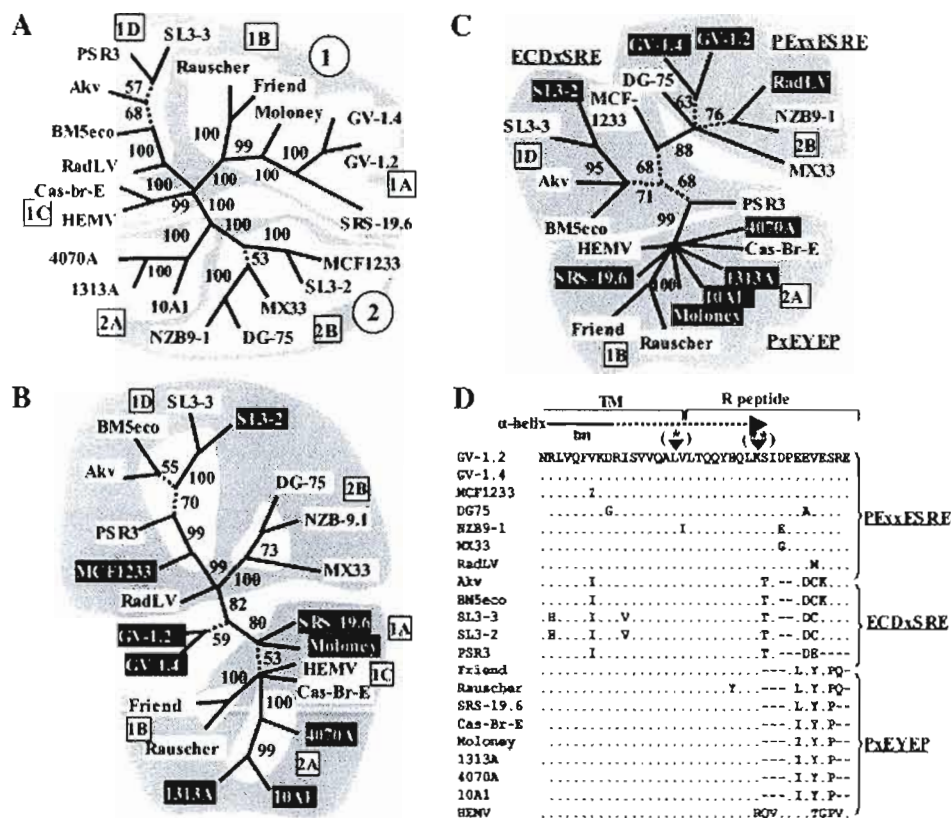


Figure 3.4: Sequential phylogenetic analysis of the Env gene. The phylogenetic trees corresponding to the SU protein (A), the TM protein and R peptide (B) and the last 150 nucleotides of TM comprising the R peptide (C). Bootstrap values are indicated for each branch of the trees. When the bootstrap value is less than 85%, the branch is drawn with a dashed line. The MuLVs represented in black boxes indicate those that changed position in the tree in comparison with the ENV tree (figure 3.2D). For the SU tree (A), the two major groups are indicated in dashed grey and numerated (1 and 2) but this was not possible for the two other trees (B and C). The subgroups are also numerated (1A-D to 2A-B). (D) Amino acid alignment of the R peptide of several MuLVs. The three major groups with consensus motifs PExxESRE, xECDxSRE and PxEYEPx are illustrated as in panel C. tm: transmembrane domain of TM. (*): protease cleavage site, (**): end of the α -helix

Figure 3.5

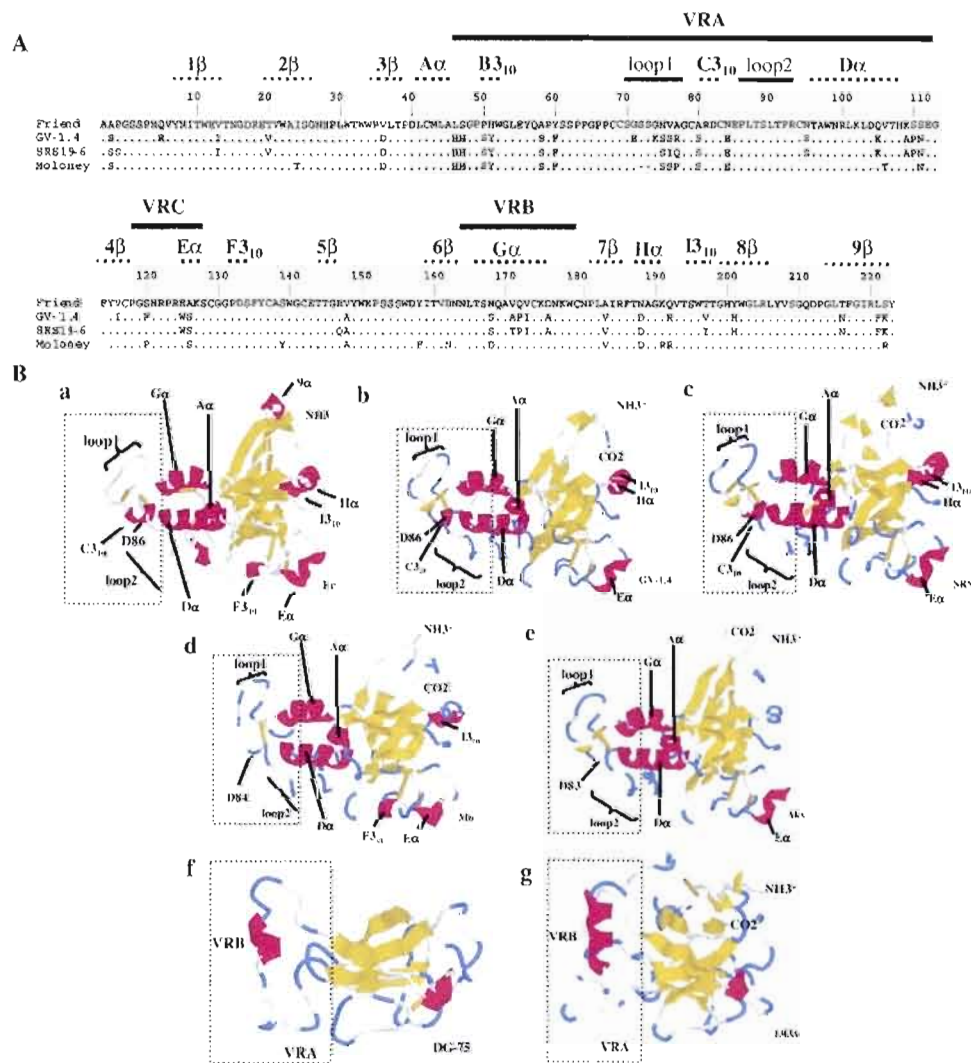


Figure 3.5: Region Binding Domain of SU (A) Multi-Alignment of the amino-acid sequence of the RBD of Friend, GV-1.4, SRS 19-6 and Moloney MuLV ENV genes. Only the non-conserved amino-acids are shown. The secondary structure elements of the Friend RBD (1aol.pdb) are indicated above the amino acid alignment. (B) Representation of homology modelled RBDs of ecotropic, amphotropic and xenotropic MuLVs. The structures were computer derived from the reference Friend RBD structure (a). The graphical representations are similarly oriented. The rectangle box on each graph indicates the receptor binding pocket. (a: Friend-MuLV, b: GV-1.4, c: SRS 19-6, d: Moloney MuLV, e: Akv, f: 1313A, g: DG-75).

Figure 3.6

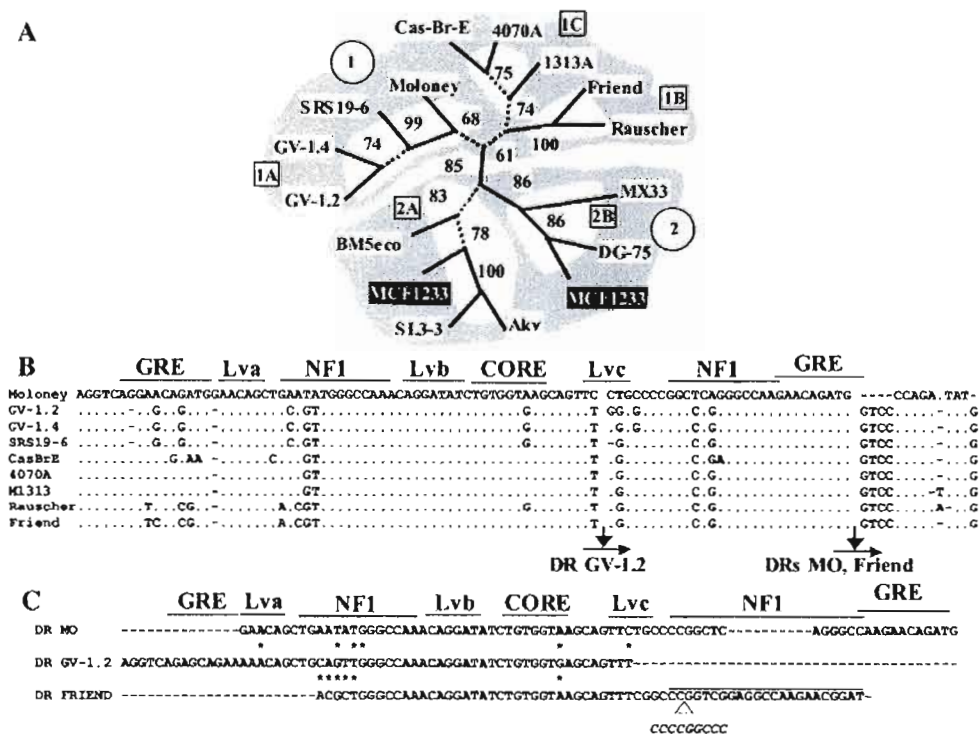


Figure 3.6: Phylogenetic analysis and multiple alignment of the LTR region. (A) The phylogenetic tree corresponding to the LTR region. When the bootstrap value was lower to 85%, the branch was drawn with a dashed line. The MuLVs represented in black boxes indicated those that changed position in the tree in comparison with the whole genome tree (figure 3.2A). For each tree, the two major groups are indicated in dashed grey and numerated (1 and 2). The subgroups are also numerated (1A-D to 2A-B). (B) Comparison of the U3 enhancer region of several MuLVs. The dots indicate similar nucleotides compared to the Mo-MuLV sequence and the dash indicate nucleotide deletion. Binding sites for transcription factors from Moloney and Friend MuLVs are boxed above and under the sequence respectively. To conserve the alignments, the direct repeats of GV-1.2 (DR GV-1.2), Mo (DR MO) and Friend (DR Friend) are indicated below the sequence with an arrow indicating the position. (C) Comparison of the Mo, Friend and GV-1.2 direct repeats. Differences are indicated by a *.

Table 3.1

Table 1: list of the MuLVs used in this study			
name	source of isolation	accession #	references
EXOGENOUS ECOTROPIC			
Graffi GV-1.2	SarcomaI, SarcomaII, Ehrlich ascites	AB187565	Graffi et al. 1957, Ru et al. 1993, Voisin et al. 2006
Graffi GV-1.4	SarcomaI, SarcomaII, Ehrlich ascites	AB187566	Graffi et al. 1957, Ru et al. 1993, Voisin et al. 2006
Cas-Br-E	wild mouse, Lake Casitas, California, USA	X57540	Gardner 1978, Perryman et al. 1991b
Friend	Ehrlich ascites carcinoma	Z11128	Friend 1957, Perryman et al. 1991a
Moloney	Sarcoma 37	J02255	Moloney 1960, Shinnick et al. 1981
Rauscher	transplantable mouse lymphoblastoma inoculated in BALB/c mice	U94692	Rauscher 1962, Khimani et al. 1997
RadLV	radiation of C57BL/6 and serial passages into C57BL/Ka	K03363	Rassart 1986, Meeregaert et al. 1987
SL3-3	tumor cell derived from the AKR mouse strain	AF169256	Pedersen et al. 1981, Lund and Pedersen 1999
SRS-19.6	TSZ (Tianjin-Shanghai-Zhejiang) system, People's Republic of China	AF019230	Bundy et al. 1995, Bundy and Fan 1998
AMPHOTROPIC			
1313 A	wild mouse, Hartz Mountain bird seed plant, California, USA	AF411814	Gardner 1978
4070A(*)	wild mouse, Lake casitas, California, USA	M33469	Gardner 1978, Ott et al. 1990
MODIFIED AMPHOTROPIC			
10A1(*)	mouse injected with 1504A, (wild mouse, La puente, California) isolate	M33470	Gardner 1978, Ott et al. 1990
XENOTROPIC			
DC-75	human B-lymphoblastoid cell line DC-75	AF221065	Raish et al. 2003
NZB9-1(*)	DNA from NZB-TU-6-infected minl: cells	K02730	O'Neill et al. 1985
ENDOGENOUS ECOTROPIC			
Akv	AKR mice	J01998	Gross 1961, 1967, Herr 1984, Erzerodt et al. 1984
HEMV	<i>Mus spicilegus</i>	AY818896	Tipper et al. 2005
PSR3	ovaries of an SWR/JxRF/J hybrid mouse	M87550	Nouvel et al. 1993
BM5eco(*)	C57BL/6 mice	AY252102	Gaur and Green 2003
POLYTROPIC			
MCF1233	isolated from C57BL mice containing Alvy-type viruses	U13766	Zijlstra et al. 1983, Sijts et al. 1994
SL3-2(*)	tumor cell derived from the AKR mouse strain	AY438266	Pedersen et al. 1981, Bahrami et al. 2004
MODIFIED POLYTROPIC			
MX33 (*)	strain HRS/J, newborn brain DNA	M17327	Stoye and Coffin 1987

(*) these viruses were used only for the analysis of the ENV region

Figure 3.7

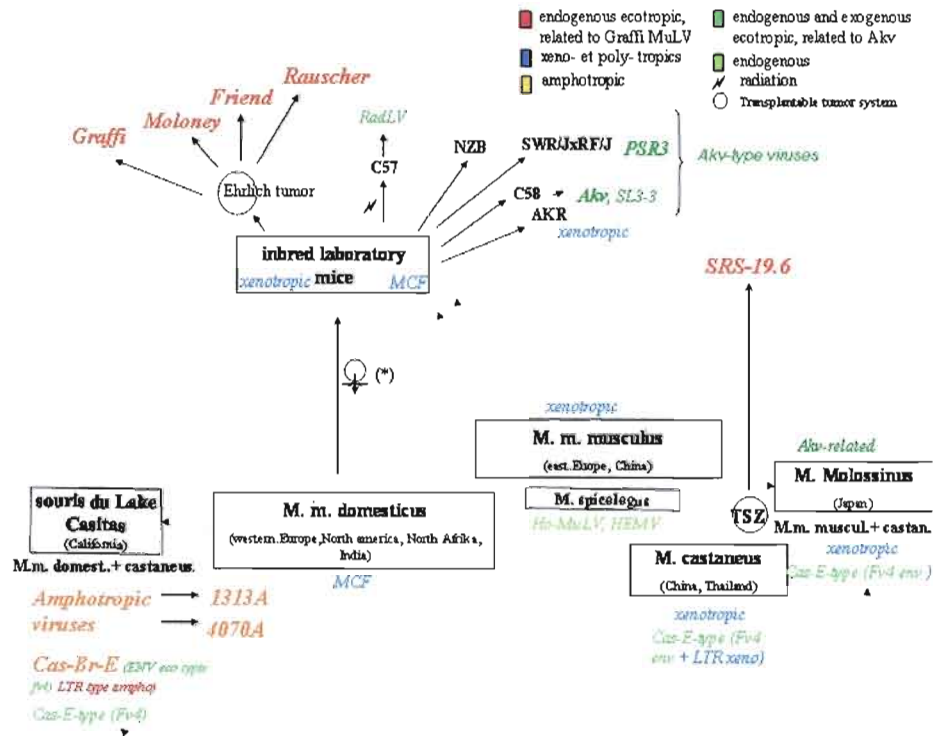


Figure 3.7: Distribution géographique des MuLVs et des souches de souris. Les MuLVs sont notés en couleur (rouge, bleu, vert, vert clair orange). Le code couleur regroupe les MuLVs les plus proches d'un point de vue phylogénique. Les souches de souris sont notées en noir et placées sur la carte du monde selon leurs origines. Cette figure est adaptée d'un article de Kozak et al., 1987, et actualisée grâce aux résultats de phylogénie.

CHAPITRE IV

PROFIL D'EXPRESSION GÉNIQUE DES LEUCÉMIES INDUITES PAR LE RETROVIRUS GRAFFI

4.1. Introduction

4.1.1. Stratégie générale

Lorsque le projet d'étudier le profil d'expression génique des leucémies induites par le rétrovirus Graffi a débuté, plusieurs méthodes étaient disponibles (*Serial Analysis of Gene Expression* (SAGE), *differential display*, *spotted* ou *macro-arrays*) et la technologie des micropuces développée par Affymetrix (Affymetrix GeneChip®). La technologie Affymetrix, bien que très coûteuse, est sans aucun doute la méthode de choix car elle est la plus globale. La micropuce GeneChip® Mouse Genome 430.2.0 choisie pour nos expériences a été assemblée en 2002 et a été conçue pour représenter l'ensemble du transcriptome murin tel que connu lors de la conception de la puce. Elle est la plus complète de toutes les micropuces du génome de la souris et contient 45000 *probesets*.

Notre objectif est de caractériser de façon approfondie les leucémies induites par le rétrovirus murin Graffi pour mieux l'utiliser comme outil de recherche sur le cancer. Nous espérons également mieux comprendre le cancer par l'analyse même de ces leucémies. Cela est en continuité avec les recherches sur les sites communs d'intégration rétrovirale qui se sont effectuées depuis de nombreuses années sur différentes tumeurs dans le laboratoire d'Eric Rassart (Denicourt et al. 1999, 2003) et qui visent directement à trouver des oncogènes. Les oncogènes qui sont ciblés régulièrement par des sites d'intégration rétrovirale ont certainement un rôle primordial dans l'apparition du cancer

et on peut les appeler premiers intervenants (ou facteur déclenchant). Mais combien de gènes 'doivent' être dérégulés pour qu'une cellule soit appelée cancéreuse? En effet, beaucoup d'autres gènes à part les gènes 'premiers intervenants' font partie de la signature d'une cellule cancéreuse et c'est ensemble qu'ils caractérisent cette cellule. Une méthode globale comme celle des micropuces Affymetrix® permet donc d'identifier la signature moléculaire des leucémies et permet de ne pas se limiter aux gènes premiers intervenants.

Le principal avantage du modèle est que le rétrovirus Graffi donne de multiples types de leucémies ayant toutes un équivalent chez l'humain. De plus, il s'agit d'un modèle murin, et donc, l'utilisation de souris congéniques atténuent fortement la déviation expérimentale. Avec une portée de souriceaux injectés le même jour avec le même échantillon viral, plusieurs types de leucémies peuvent être analysées. On minimise ainsi les variations expérimentales et les différents types de leucémies résultent uniquement du hasard des intégrations rétrovirales. Ainsi, ce modèle est idéal pour comparer les diverses leucémies entre elles et doit donc être exploité dans ce sens plutôt que d'analyser les leucémies de façon individuelle. Comparer les divers types de leucémies permet de faire ressortir ce qui est spécifique à chaque leucémie notamment les oncogènes et les voies de transmission de signaux spécifiquement employées. Ceci peut aller dans la direction de chercher de nouveaux marqueurs de diagnostic ou de nouvelles cibles thérapeutiques si on l'applique à l'humain.

4.1.2. Présentation d'une micropuce

Une micropuce Affymetrix est une pièce de silice sur laquelle sont placées des milliers de séquences d'ADN correspondant à l'ensemble des transcrits découverts chez la souris au moment de la construction de la micropuce (figure 4.1). Les séquences présentes sur la micropuce GeneChip® Mouse Genome 430. 2.0 ont été sélectionnées en 2002 à partir de l'analyse de base de données comportant des informations sur le transcriptome (GenBank®, dbEST, RefSeq, UniGene (version 107, Juin 2002) ainsi qu'en s'appuyant sur le premier assemblage du génome de la souris (Whitehead Institute Center for Genome Research, MGSC, Avril 2002 (fiche technique disponible sur www.Affymetrix.com)).

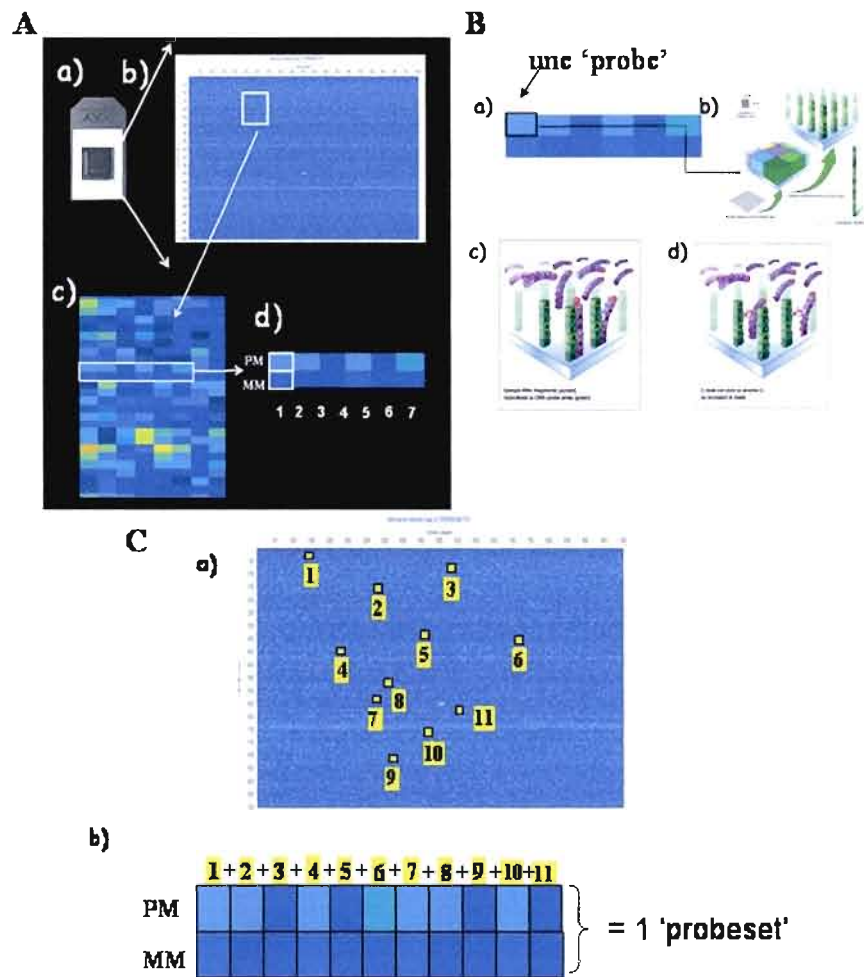


Figure 4.1: Présentation d'une micropuce Affymetrix. (A) (a) photographie d'une micropuce Affymetrix. (b) image du scanner d'une micropuce après une hybridation (c) zoom d'une partie de l'image. Chaque carré correspond à une 'probe'. (d) représentation de 7 paires de *probes* avec les carrés du haut représentant les *perfect matches* (PM) et ceux du bas les *mismatches* (MM). (B) Chaque *probe* est composée de milliers de séquences oligonucléotidiques semblables sur lesquelles vont s'hybrider les molécules d'ARNc biotinylés. (C) Sur la micropuce, un transcrite est représenté par un *probeset* qui est composé de 11 paires de *probes* (PM +MM) dispersées en plusieurs endroits sur la puce.

Les ARNs messagers totaux sont extraits à partir des cellules. Ils sont successivement transformés en molécules d'ADN complémentaires (rétrotranscription à partir d'un oligodT pour la synthèse du premier brin et polymérisation grâce à des oligonucléotides dégénérés pour la synthèse du second brin). Ensuite, une transcription *in vitro* avec des nucléosides biotinylés produit des molécules d'ARN biotinylées (ARNc) qui vont être fragmentées et vont s'hybrider aux séquences d'ADN complémentaires sur les micropuces (figure 4.2). Les molécules d'ARNc biotinylées vont être couplées à un anticorps fluorescent (anticorps streptavidine-phycoerythrine). Les micropuces vont être scannées et l'intensité de fluorescence est corrélée au nombre de copies de molécules d'ARNc hybridées aux séquences d'ADN (figures 4.1, 4.2). La micropuce permet donc de mesurer le taux relatif d'expression de chaque gène pour chaque leucémie.

Chaque séquence d'ADN est un oligonucléotide long d'environ 20 bases. Chaque transcrit est représenté sur la puce par 11 paires d'oligonucléotides. Chaque oligonucléotide est appelé une *probe*. Une paire de *probes* est constituée par un *perfect match* et un *mismatch*. Le *perfect match* est constitué d'une séquence 100% homologue au gène d'intérêt alors que le *mismatch* présente une mutation dans la séquence. L'ensemble des 11 paires de *probes* constitue un *probeset* (figure 4.1.C). Un gène peut être représenté par plusieurs *probesets* qui correspondent à différentes régions du gène. Les différents *probesets* peuvent notamment servir à différencier plusieurs transcrits alternatifs. Les 11 paires de *probes* d'un *probeset* sont dispersées sur la puce pour minimiser l'impact d'artefacts d'hybridation sur les résultats (bulles d'air par exemple). Pour analyser les résultats, il faut reconstituer une valeur unique pour le *probeset* à partir des intensités des 11 *probes*. Cette étape est nommée le calcul du sommaire d'expression. Le *mismatch* est présent pour différencier les hybridations non spécifiques des hybridations spécifiques mais son implication dans le calcul du sommaire d'expression est maintenant très controversée (section 4.3.2).

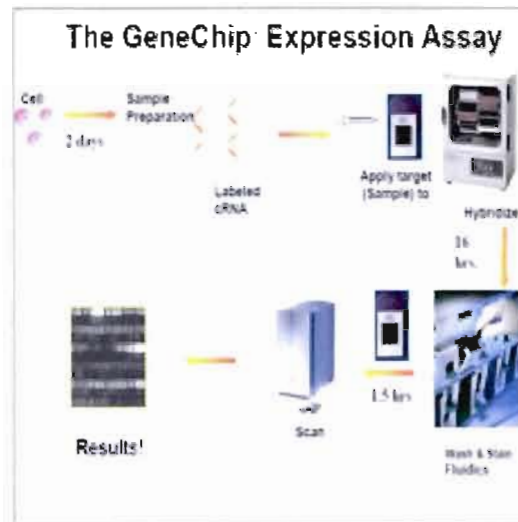
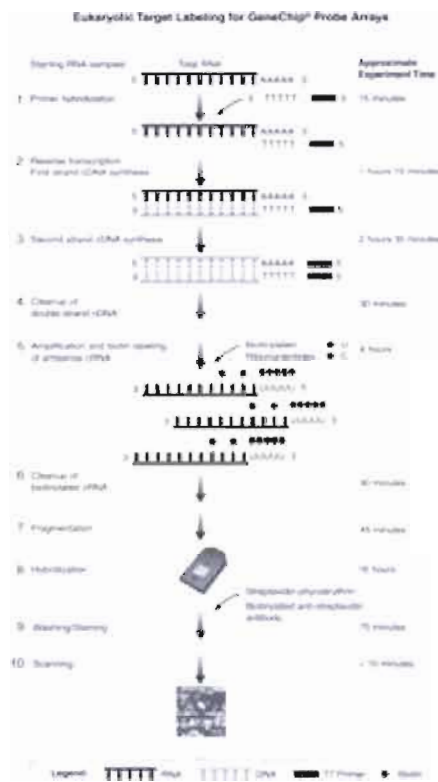
A**B**

Figure 4.2: (A) Schématisation de la procédure d’Affymetrix. (B) Schématisation du protocole pour passer de l’ARN messager à l’ARNc biotinylé, fragmenté et hybridé.

4.1.3. Tri des leucémies

Afin d'obtenir des résultats valides avec les micropuces, il a fallu au préalable trier les population leucémiques à partir des organes hématopoïétiques des souris leucémiques. Le tri cellulaire s'est effectué grâce à un système d'aimant-billes magnétiques développé par la compagnie StemCell technologies (EasySep®, www.stemcell.com). Dans ce système, les cellules leucémiques sont couplées aux billes magnétiques grâce à des complexes d'anticorps. Les cellules leucémiques sont alors retenues par l'aimant et triées par sélection positive. La sélection négative a été testée mais n'a pas conduit à une pureté suffisante et n'a donc pas été retenue pour les échantillons analysés par micropuce. Les critères de qualité pour le tri ont été fixés à un minimum de 95% de pureté et un minimum de 95% de viabilité après le tri pour les échantillons envoyés aux micropuces.

Après le tri cellulaire, les ARNs messagers totaux ont été extraits au Trizol® à partir des cellules triées puis purifiés par passage sur des colonnes (RNeasy, QIAGEN®). Les échantillons ont alors été confiés au centre génomique de McGill (Centre d'innovation Génome Québec et Université McGill, Montréal) afin de procéder à la transformation en ARNc biotinylés et à d'hybridation sur les micropuces GeneChip 430.2®.

4.2. Phénotypes des leucémies triées et envoyées à l'analyse par micropuces

Seize échantillons ont été envoyés pour l'analyse du profil d'expression génique dont 13 correspondent à des leucémies et 3 à des contrôles. Alors que pour les leucémies, chaque échantillon provient d'une souris, les cellules provenant de 12 souris ont été mises ensemble pour constituer chaque échantillon contrôle (6 mâles et 6 femelles). Les phénotypes leucémiques ont été choisis pour que cela soit un reflet de l'éventail des différentes leucémies induites par le rétrovirus Graffi (chapitre II). Les cellules des leucémies de type T ont été triées à partir de thymus (thymomes), les cellules leucémiques de type B à partir des ganglions, et les autres types (myéloïde, érythroïdes et

mégacaryocytaires) à partir de la rate. Le tableau ci-dessous (tableau 4.1) présente les phénotypes des différents échantillons envoyés aux micropuces.

Tableau 4.1: leucémies hybridées sur les micropuces

phénotypes:	micropuce		triée avec:	organe:
CD4 ⁺ CD8 ⁺	T1	MOT-01	anti-CD4	thymus
CD4 ⁺ CD8 ⁻	T2	MOT-02	anti-CD4	thymus
CD4 ⁻ CD8 ⁺	T3	MOT-03	anti-CD8	thymus
B220 ⁺ CD19 ⁺	B1	MOT-04	anti-CD19	ganglions
B220 ⁺ CD19 ⁺	B2	MOT-05	anti-CD19	ganglions
B220 ^{low} CD19 ⁺ Sca-1 ⁺	B3	MOT-06	anti-CD19	ganglions
CD11b ⁺ Gr-1 ⁺	M	MOT-07	anti-CD11b	rate**
Ter119 ⁺ CD71 ⁺	E1	MOT-08	anti-Ter119	rate**
Ter119 ⁺ CD71 ⁺	E2	MOT-09	anti-Ter119	rate**
Ter119 ⁺ CD71 ⁺	E3	MOT-10	anti-Ter119	rate
cKit ⁺ CD41 ⁻	Mk1	MOT-11	anti-cKit	rate
cKit ⁺ CD41 ⁺	Mk2	MOT-12	anti-CD41	rate
cKit ⁺ CD41 ⁺	Mk3	MOT-13	anti-CD41	rate
échantillons contrôles:				
CD4 ⁺ CD8 ⁺ (⁺)	CT	MOT-14	anti-CD4, -CD8	thymus
B220 ⁺ CD19 ⁺ (⁺)	CB	MOT-15	anti-CD19	moelle osseuse
Ter119 ⁺ CD71 ⁺ (⁺)	CE	MOT-16	anti-Ter119	moelle osseuse**
*n=12; ** lyse des globules rouges avant le tri				

L'hybridation des micropuces suit un protocole très standardisé et est effectuée par le même manipulateur. Ainsi, il n'est pas nécessaire d'effectuer plusieurs puces avec le même échantillon d'ARN (tests effectués par l'équipe du centre génomique de McGill, A. Ponton). La plus petite quantité d'ARN total a été obtenue pour les leucémies mégacaryocytaires (échantillons Mk1 et Mk2), ces leucémies étant moins invasives (résultats supplémentaires, chapitre II). La quantité d'ARN total utilisée pour l'hybridation des micropuces a été fixée à 2 µg pour tous les échantillons. Pour les leucémies de type myéloïde, il n'a pas été possible d'obtenir 3 échantillons de haute qualité (population non pure et leucémies mixtes). Ainsi une seule leucémie myéloïde est incluse dans l'analyse par micropuces. Bien qu'il n'y ait aucun problème au niveau de la normalisation de la micropuce (voir plus bas), la leucémie myéloïde ne pourra pas être incluse dans des tests statistiques de type Student, le nombre de réplifications minimum étant de 3. Ainsi l'échantillon est présent à titre indicatif et si des informations semblent

provenir de cet échantillon au cours de l'analyse, des preuves supplémentaires doivent nécessairement être apportées par un échantillonnage plus grand de leucémies myéloïdes.

4.3. Résultats Partie I: validation de l'expérience

Les analyses et graphiques de cette section ont été réalisés avec le logiciel FlexArray (beta 3.1, version 0.7.1.0) développé par Michal Blazejczyk et Mathieu Miron du centre génomique McGill (Montréal).

4.3.1. Présentation des données brutes

Les données brutes correspondent aux valeurs attribuées à chaque *probe* à l'issue du scanner des micropuces (figure 4.3). Chaque valeur est corrélée à l'intensité de fluorescence et donc relative au nombre de molécules d'ARNc qui se sont hybridées. L'observation des images scannées (figure 4.3) qui est la première étape de validation de l'expérience a permis de vérifier qu'il n'y avait pas d'artéfacts majeurs d'hybridation des micropuces. Deuxièmement, l'observation du graphique de répartition des intensités (figure 4.4.A) des micropuces non normalisées permet d'apprécier la répartition de l'intensité globale de chaque micropuce et d'écarter de possibles échantillons à problème. Ainsi on peut observer que CP050802MOT-11 (échantillon Mk1) a l'intensité globale la plus faible avant normalisation (figure 4.4.A). Malgré cela, la différence n'est pas majeure et il n'est pas nécessaire d'exclure cette micropuce de l'analyse. La figure 4.4.B montre l'effet de la normalisation sur la répartition des données (section 4.3.2).

4.3.2. Présentation des données normalisées

Les données brutes d'intensité de fluorescence provenant des micropuces scannées doivent être transformées par des méthodes mathématiques pour pouvoir comparer les données de plusieurs micropuces entre elles et calculer les valeurs de chaque *probeset*. Quand Affymetrix a commercialisé les premières puces, la normalisation se faisait essentiellement par rapport au bruit de fond général et par rapport à la valeur des gènes

housekeeping comme Gapdh ou l'actine (comme on le fait pour la quantification de résultats de RT-PCR). Le sommaire d'expression pour obtenir la valeur d'un *probeset* se faisait par la soustraction des valeurs du *perfect match* par rapport à celle du *mismatch*, puis par le calcul de la moyenne des valeurs des *probes* constituant le *probeset*.

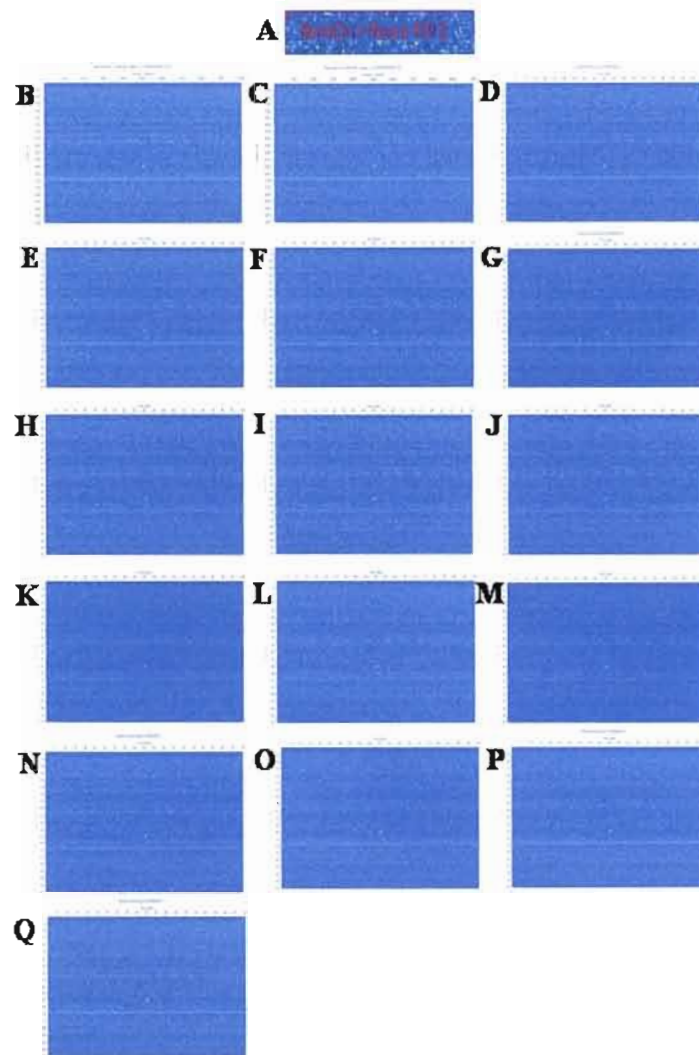


Figure 4.3: Images des données brutes des micropuces. (A) contrôle de qualité d'hybridation: titre placé en haut et à droite de chaque puce. B: T1 (CP050802MOT-01), C: T2 (MOT-02), D: T3 (MOT03) E: B1(MOT04) F: B2 (MOT05) G: B3(MOT06) H: M(MOT07) I: E1(MOT08) J: E2(MOT09) K: E3(MOT10) L: MK1(MOT11) M: MK2(MOT12) N: MK3(MOT13) O: CT(MOT14) P: CB(MOT 15) Q: CE(MOT16)

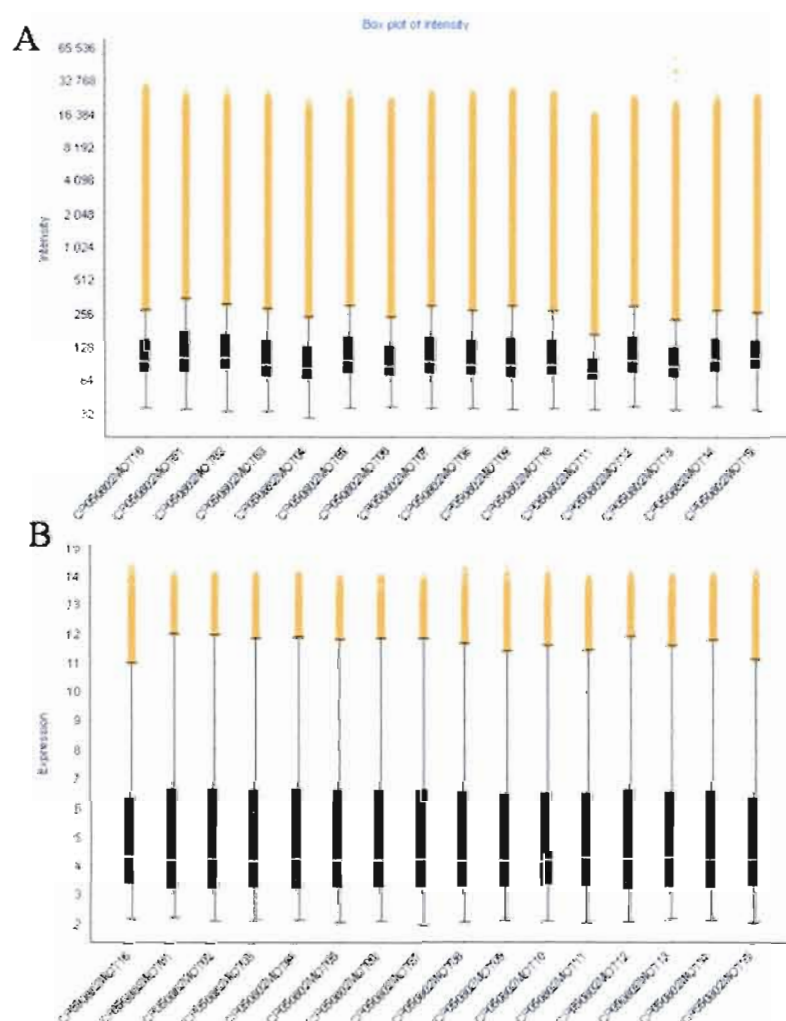


Figure 4.4: Graphiques de distribution des intensités de chaque micropuce. (A) Données brutes. **(B)** Après normalisation. Le rectangle noir représente l'échelle d'intensité de 50% des données.

Lors de tests de simulation effectués par des biostatisticiens, de nombreux biais sortant des analyses de micropuces ont été révélés, notamment de nombreux taux de faux-positifs et faux-négatifs (erreurs de type I et II respectivement). Ainsi des méthodes de plus en plus sophistiquées apparaissent continuellement pour normaliser les micropuces et calculer le sommaire d'expression. On peut citer les plus connues: d-Chip (Li and Wong, 2001), RMA (Irizarry, 2003), GC-RMA (Wu and Irizarry, 2005), MAS5.0 (Microarrays

Analysis Suite 5.0, Affymetrix, 2001). Un changement majeur a été l'application d'une échelle logarithmique aux valeurs d'intensité dans toutes les méthodes de calcul.

RMA (*Robust multi-array analysis*) a été choisie comme méthode préférentielle dans cette présente analyse. RMA est estimée être une méthode plus robuste que le plus récent produit d'Affymetrix, MAS5.0, ou même dChip (Cope et al. 2004). Un des avantages de la méthode est que les données de l'ensemble des micropuces sont utilisées pour établir la relation de normalisation alors que certaines méthodes comme MAS5.0 normalisent les résultats par rapport à une micropuce désignée. Une bonne normalisation est réalisée avec RMA en prenant un nombre minimal de micropuces égal ou supérieur à 6. De plus, RMA ne tient pas compte des valeurs du *mismatch* pour calculer le sommaire d'expression d'un *probeset* comme le fait encore MAS5.0. A l'issue de la normalisation, une valeur en log base 2 est attribuée à chaque *probeset* de chaque micropuce. Ainsi un changement en valeur logarithmique (ratio) égal à 1 entre 2 micropuces pour un *probeset* donné équivaut en réalité à un changement d'expression de 2 fois (*fold change*). Un ratio d'un minimum de 0,585 (changement de 1,5 fois) est utilisée généralement comme critère de filtre dans les analyses de micropuces avec RMA. En dessous de cette valeur les résultats ont une forte probabilité d'être de faux-positifs.

MAS5.0 correspond à la plus récente version de l'algorithme statistique développé par la compagnie Affymetrix. A partir des données brutes, la première étape de calcul est la normalisation de l'intensité des micropuces entre elles avec une correction du bruit de fond. Ensuite différentes valeurs relatives à l'expression des transcrits vont être calculées. Deux valeurs fondamentales sont issues de MAS5.0 à partir des données brutes: la valeur de détection et le signal.

La valeur du signal est le calcul du sommaire d'expression pour chaque *probeset* comme le fait RMA mais en prenant compte les valeurs du *mismatch*.

Le calcul de la valeur de détection (*call value*) est établi par rapport à la comparaison entre l'intensité du fragment non muté (*perfect match*) et celle du fragment muté (*mismatch*) pour un *probeset* donné. Si la valeur de détection est calculée *présente*, cela donne une indication que le gène correspondant est bien exprimé dans les cellules (il y a présence d'un nombre significatif de copies d'ARNm de ce gène). Le raisonnement est inverse (absence d'expression de ce gène) lorsque la valeur de détection est calculée

absente. Une valeur statistique p est associée à la valeur de détection et estime le taux de confiance que l'on peut accorder à cette valeur.

MAS5.0, bien qu'améliorée, est encore controversée à cause de l'implication du *mismatch* dans les calculs. Cependant des comparaisons entre les résultats de RMA et de MAS5.0 dans l'étude présente ont montré une très bonne corrélation entre les 2 méthodes. La valeur de détection est appliquée dans le cas présent pour confirmer les résultats de RMA. En effet, si la valeur de détection *absente* peut être un faux-négatif (un gène est effectivement présent mais il est détecté absent à cause du *mismatch*), une valeur de détection *présente* ne peut pas être un faux-positif. Ainsi les valeurs '*présentes*' nous confirment l'expression de ce gène dans les cellules (avec un nombre de copies faibles ou élevées).

La figure 4.4.B montre le résultat de la normalisation des micropuces par RMA. On peut voir que les intensités globales sont maintenant comparables, et que l'échantillon CP050802MOT-11 (échantillon Mk1) n'a plus une intensité globale faible. On peut aussi observer que les échantillons CP050802MOT-15 et CP050802MOT-16, qui correspondent aux contrôles CB et CE (tableau 4.1), sont un peu différents. Ces différences sont minimales. Des analyses avec normalisation des données avec et sans les contrôles ont montré les mêmes résultats.

4.3.3. Vérification de la validité de l'expérience

Les leucémies induites par le rétrovirus Graffi sont très hétérogènes (chapitre II) et les phénotypes choisis expressément différents. On ne pourra trouver des gènes spécifiques de chaque type de leucémie que si les leucémies à l'intérieur d'un même groupe (par exemple les leucémies de type T) sont plus semblables entre elles que par rapport aux autres types de leucémies.

La méthode PCA (Analyse en Composantes Principales) est une méthode de choix pour tester cette hypothèse (Misra et al. 2002). Le logiciel FlexArray a été utilisé pour obtenir les graphiques PCA (figure 4.5). La méthode PCA consiste à réduire la dimension des données pour visualiser clairement les tendances existantes entre les données. C'est un outil d'exploration des données qui ne fait intervenir *aucun critère présélectionné par l'utilisateur* et est donc différent des outils de *clustering*. La méthode capture de façon

linéaire le maximum de données à partir du *set* initial pour créer un *set* plus petit mais contenant les informations essentielles et discriminatoires pour comparer les données entre elles.

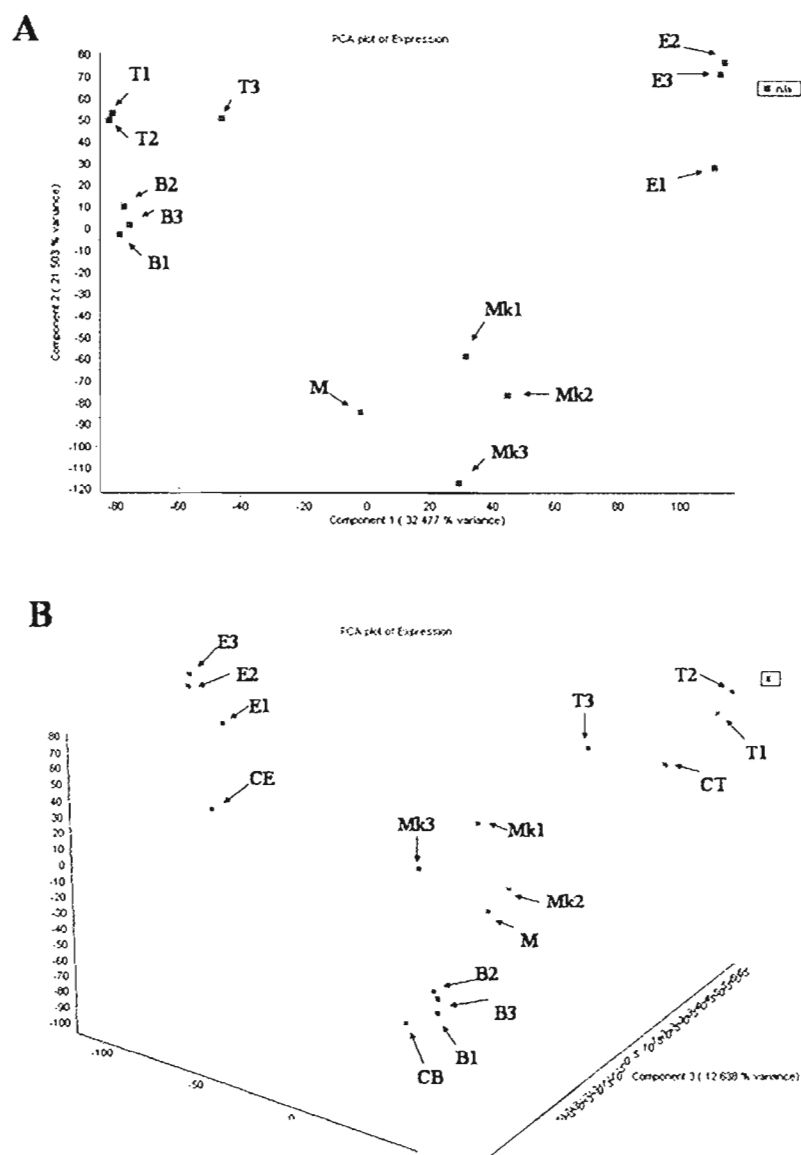


Figure 4.5: Graphiques PCA. (A) Echantillons leucémiques. (B) Echantillons leucémiques avec les contrôles.

Cette méthode donne des résultats très convaincants sur notre jeu de données (figure 4.5). La figure 4.5.A montre le graphe comprenant toutes les leucémies sans les contrôles (normalisation RMA avec les échantillons leucémiques seulement). Très clairement, chaque type de leucémies (T, B, E, Mk) se regroupe de façon individuelle. Ainsi l'hypothèse que nous allons pouvoir faire ressortir les gènes spécifiques à chaque type de leucémie à partir des résultats des micropuces, est vérifiée. Il est à noter que la leucémie myéloïde est proche des leucémies mégacaryoblastiques.

Pour utiliser les échantillons contrôles dans l'analyse, il a fallu également vérifier l'hypothèse que chaque contrôle est plus proche du type de leucémies auquel il correspond par rapport aux autres échantillons. Cette hypothèse est également vérifiée (figure 4.5.B) et les leucémies vont pouvoir être comparées par rapport à leur contrôle respectif (normalisation RMA de l'ensemble des données)

Les diagrammes de dispersion (non montré) ont également permis de comparer la répartition des données entre 2 micropuces. Il est apparu que les micropuces issues de 2 types différents de leucémie étaient plus différentes entre elles que si elles provenaient du même groupe.

En conclusion ces analyses ont démontré la validité de l'expérience et la partie suivante va consister à rechercher les spécificités de chaque type de leucémies.

4.4 Résultats Partie II: Profil général des leucémies

A partir des données RMA des échantillons leucémiques, une analyse de *clustering* a été réalisée avec une sélection de gènes montrant les plus grandes variations entre les leucémies. Le but de cette analyse est de *découvrir* comment se répartissent les gènes en fonction des types de leucémies.

4.4.1 Analyse différentielle

L'unique possibilité pour traiter les données issues des résultats de micropuces est de faire des analyses différentielles, c'est-à-dire de comparer l'intensité d'un *probeset* entre différentes micropuces. Il n'est pas possible de comparer l'intensité de 2 *probesets* dans

la même micropuce. Par exemple, il est possible de dire que *c-myc* est exprimé plus fortement dans T1 par rapport à T2 mais il est impossible de dire que *c-myc* est exprimé plus fortement que *Notch1* dans T1. Ceci est dû au fait que l'efficacité d'hybridation diffère d'une *probe* à l'autre. Par exemple, plus la séquence d'une *probe* est riche en G+C et plus la stabilité de l'hybridation est grande (Kluger et al. 2003).

Il est possible de faire ressortir de l'analyse les gènes spécifiques à chaque type de leucémie et d'observer les expressions de chaque gène par rapport à la moyenne. Ainsi la valeur moyenne d'expression pour chaque *probeset* a été calculée à partir des données de toutes les micropuces. Ensuite, le ratio a été calculé pour chaque leucémie par rapport à la moyenne. Une valeur positive de ratio signifie une surexpression du *probeset* dans l'échantillon leucémique par rapport à la moyenne des leucémies et une valeur négative signifie une sous-expression. Un exemple de calcul est montré par le tableau 4.2.

Tableau 4.2: exemple de calcul des ratios

1421854_at	T1	T2	T3	B1	B2	B3	M	E1	E2	E3	Mk1	Mk2	Mk3	moyenne
données RMA	4,81	2,75	6,32	5,20	4,93	5,01	8,50	4,21	4,21	3,52	4,57	6,01	5,60	5,05
ratios*	-0,24	-2,30	1,27	0,15	-0,12	-0,05	3,45	-0,84	-0,84	-1,52	-0,48	0,96	0,55	

*ratio=moyenne-données RMA

Dans l'exemple ci-dessus, la moyenne d'expression du *probeset* 1421854_at est de 5,05. Après le calcul des ratios, on voit que les échantillons T3, B1, M, Mk2 et Mk3 montrent une surexpression du *probeset* par rapport à la moyenne alors que les autres échantillons montrent une sous-expression. Un ratio inférieur à la valeur absolue 0,585 n'est pas considéré comme étant significatif (haute probabilité d'être un faux-positif). La couleur rouge correspond à une surexpression supérieure ou égale à 0,585 par rapport à la moyenne et une la couleur correspond à une sous-expression inférieure ou égale à 0,585. Les valeurs sont en logarithme de base 2, et un ratio de 3,45 (échantillon M) équivaut à un changement de 10,9 fois entre l'échantillon leucémique et la moyenne. Cette technique de calcul a été pour la première fois utilisée par l'équipe de T. Golub qui fut le premier à établir des sous-classes de leucémies et à les relier à différentes chances de guérison (Golub et al. 1999).

4.4.2. Sélection de 3000 gènes

Les micropuces permettent d'évaluer chacune l'intensité de 45000 *probesets*. Pour l'analyse par *clustering* réalisée avec le logiciel GeneCluster (voir plus bas), il a fallu sélectionner 3000 gènes, ceci étant dû aux limitations de l'algorithme. De plus, il est logique de débiter un projet de recherche en se limitant dans le nombre de données et de commencer par celles qui sont à coup sûr très significatives.

Les gènes les plus intéressants sont ceux qui présentent les plus grandes variations entre les leucémies et donc dont la somme des ratios est la plus grande. Les *probesets* ont alors été ordonnés de façon décroissante par rapport aux valeurs des ratios et les 3000 premiers *probesets* ont été sélectionnés.

Parmi les gènes qui ne montrent pas d'expression différentielle entre les leucémies, il se trouve des gènes qui sont exprimés à des niveaux équivalents dans toutes les cellules comme les gènes *housekeeping* par exemple mais aussi de nombreux gènes qui ne sont pas exprimés dans les cellules de type hématopoïétique. En effet, les *probes* disposées sur les micropuces d'Affymetrix représentent l'ensemble des gènes qui sont connus pour être exprimés dans l'organisme murin mais le transcriptome diffère selon le type cellulaire.

4.4.3. Clustering hiérarchique, analyse SOMs et profils des leucémies

4.4.3.1 Principes des analyses de *clustering*

Le but de ces analyses est de découvrir la répartition des données pour chaque type de leucémie et cela va permettre de savoir sur quels patrons d'expression l'attention doit être portée. *Est-ce que les résultats sont tels qu'attendus ou est-ce que certains gènes se regroupent de façon inattendue?*

Le *clustering* hiérarchique a été adapté au traitement des données de micropuces par Michael Eisen (Eisen et al. 1998, Chu et al. 1998). Le processus du *cluster* hiérarchique consiste à assembler un ensemble d'items (*probesets* ou micropuces) sous forme d'un arbre dont les items qui se ressemblent le plus vont être reliés par des branches très courtes alors que les items qui se ressemblent le moins vont l'être par des branches très longues. La première étape de la construction de l'arbre par l'algorithme est la

construction d'une matrice de distance. Ensuite les 2 items dont la distance est la moins grande sont regroupés pour former un pseudo-item (*cluster*) et le regroupement entre les items continue ainsi de suite jusqu'à ce qu'il ne reste plus de données à joindre.

Les analyses SOMs (*Self-Organization Maps*) réarrangent également les gènes ou les micropuces selon la similarité des valeurs d'intensité. La différence majeure est que le nombre de clusters est fixé à l'avance par l'utilisateur. De plus, les clusters sont organisés de telle façon que les clusters voisins ont entre eux des similarités. Les analyses SOMs ont été développées par T. Kohonen (Kohonen 1999) et ont aussi été utilisées par T. Golub pour les analyses de micropuces (Golub et al. 1999).

4.4.3.2. Profil général des leucémies

Un cluster hiérarchique a été réalisé sur les 3000 gènes sélectionnés, grâce au programme *Gene Cluster* développé par M. Eisen (<http://rana.stanford.edu/software>) et visualisé grâce au programme *Treeview* (<http://rana.stanford.edu/software>). Le résultat est présenté sur la figure 4.6.A. Chaque ligne correspond à un *probeset* et chaque colonne correspond à une micropuce. Une couleur verte indique une valeur négative de ratios (donc une sous-expression) et une couleur rouge indique une valeur positive (surexpression) par rapport à la moyenne des leucémies. Plus les couleurs verte ou rouge sont intenses et plus la valeur absolue du ratio est grande.

Le dendrogramme à gauche (figure 4.6.A) représente comment les *probesets* se regroupent entre eux et le dendrogramme du haut représente les similarités entre les micropuces. On voit que les *probesets* ainsi que les micropuces se regroupent bien selon les types de leucémies. Par exemple, le nœud noté X sur la figure 4.6.A regroupe bien les *probesets* qui ont tendance à être surexprimés dans les leucémies non-lymphoïdes et sous-exprimés dans les leucémies lymphoïdes par rapport à la moyenne.

L'agrandissement du dendrogramme des micropuces (figure 4.6.C) montre que les leucémies lymphoïdes et non-lymphoïdes forment 2 grands groupes distincts. Au sein du groupe des lymphoïdes, les leucémies de type T se regroupent entre elles et celles de type B aussi. Au sein des leucémies non-lymphoïdes, les érythroïdes forment un groupe distinct des échantillons mégacaryocytaires. La leucémie de type myéloïde est plus proche des leucémies mégacaryoblastiques. Ceci est similaire aux résultats obtenus avec

l'analyse en composantes principales (section 4.3.3.1). Ainsi de nouveau, cela permet de valider l'expérience et indique qu'il est possible de faire ressortir les gènes spécifiques de chaque type de leucémies à partir de ces données. De plus, cela confirme que le calcul des ratios est une bonne méthode pour accéder à ce but.

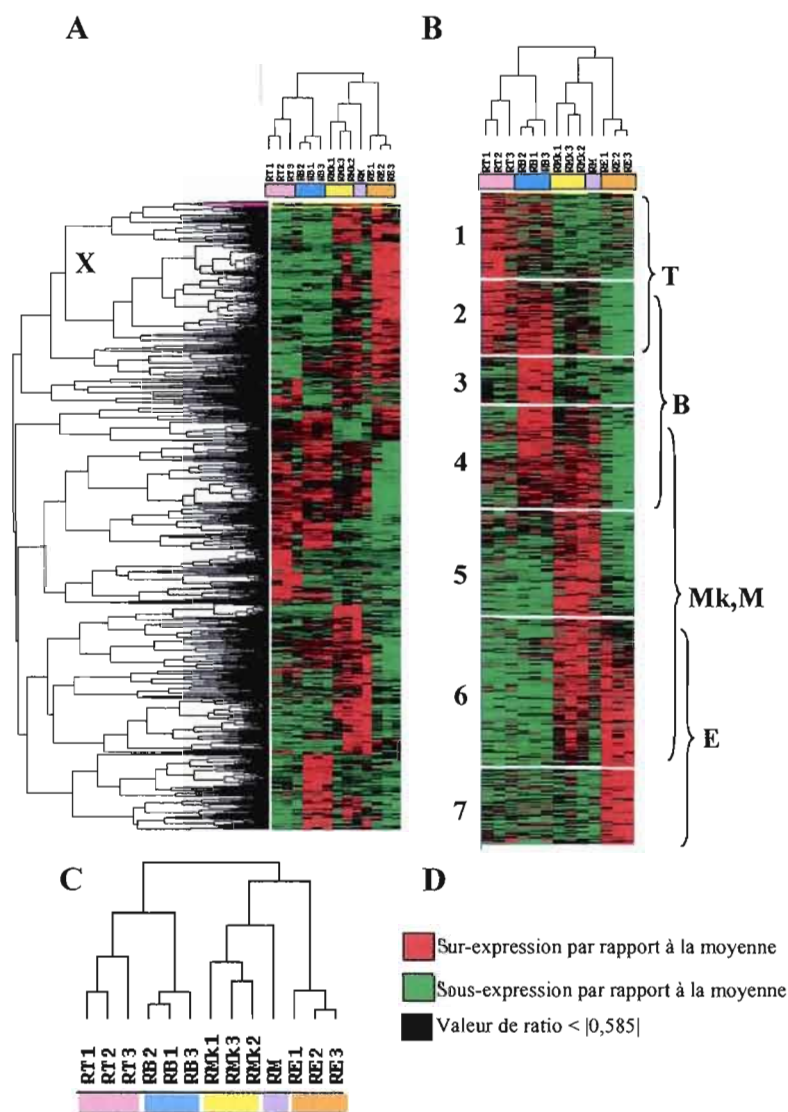


Figure 4.6: Visualisation des clusters hiérarchiques. (A) Résultat du cluster sur les 3000 gènes et dendrogramme du cluster des gènes. (B) Résultat du cluster hiérarchique après l'analyse SOM. (C) Zoom sur le dendrogramme correspondant au *clustering* des micropuces. (D) Présentation du code couleur pour les valeurs de ratios.

Pour obtenir un profil général de meilleure qualité, une analyse SOM (GeneCluster, paramètres G 1-5; A 1-10) a été combinée avec un *cluster* hiérarchique (paramètres *complete linkage clustering, correlation (uncentered), cluster on arrays only*) pour donner un résultat convaincant présenté en figure 4.6.B. Quatre grands groupes de gènes de taille semblable sont formés par la similarité de leurs patrons d'expression et se chevauchent. Le groupe 1 correspond aux gènes plus spécifiques des leucémies de type T, le groupe 2 aux leucémies de type B, le groupe 3 aux leucémies de type mégacaryoblastique et à la leucémie myéloïde et le groupe 4 aux leucémies de type érythroïde. Le cluster fait bien la distinction entre leucémies lymphoïdes et leucémies non-lymphoïdes.

4.4.3.3 Profil de chaque type de leucémies et conclusions

Le profil de chaque type de leucémie a été étudié plus en détail. Afin d'établir chacun de ces profils, les gènes qui sont *surexprimés dans les 3 leucémies d'un même type* par rapport à la moyenne sont sélectionnés. Une analyse SOM est alors réalisée sur les gènes correspondant et visualisée par *Treeview* (figure 4.7). Les sous-groupes formés par l'analyse SOM sont séparés par des traits bleus et numérotés. Chacune de ces figures représente la signature génique de ces leucémies. Elles représentent une vue d'ensemble des gènes exprimés dans chaque type de leucémies. Les leucémies ont entre elles des signatures communes ou divergentes qu'il est utile de connaître. Les signatures communes se retrouvent sur plusieurs profils, par exemple les gènes surexprimés communément dans les leucémies T et B sont représentés par les groupes C2 et D3.

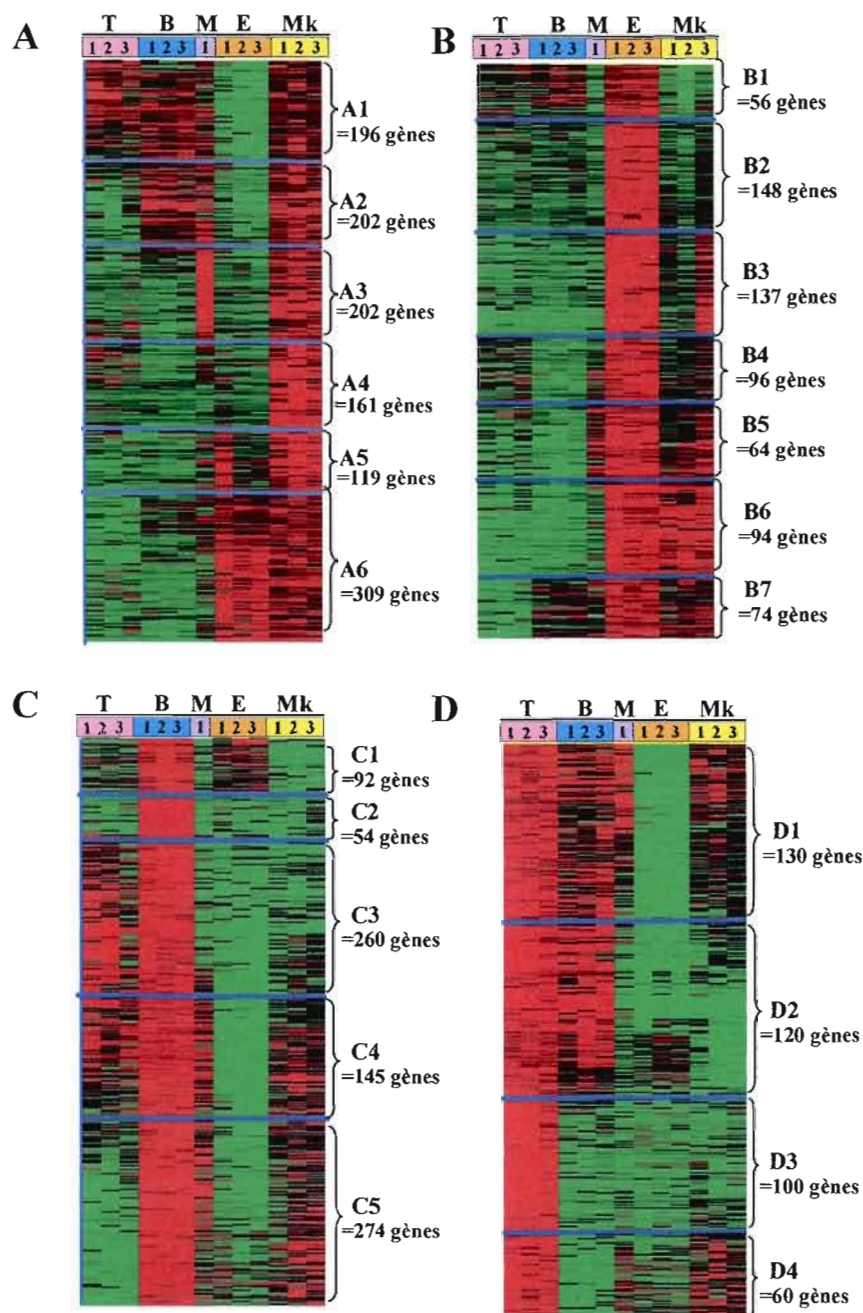


Figure 4.7: Profil général des leucémies (analyses SOMs visualisées avec Treeview). (A) Gènes surexprimés dans les leucémies mégacaryoblastiques. (B) Gènes surexprimés dans les leucémies érythroïdes. (C) Gènes surexprimés dans les leucémies B. (D) Gènes surexprimés dans les leucémies T. Les groupes formés par SOM sont séparés par des traits bleus, numérotés et le nombre de gènes impliqués est également indiqué.

Le tableau 4.3 résume les résultats obtenus à partir de ces 4 profils et regroupe les signatures communes.

Tableau 4.3: Groupes identifiés par les analyses SOM

	description	micropuces					numéros des groupes (figure 4.7)
		T	B	M	E	Mk	
1	T						D3
2	B						C2
3	T+ B						C3 , D2
4	Mk						A4
5	E						B2
6	Mk+ E						A6, B6
7	M+ Mk+ E						B5 (+-), A6(+)
8	T+ B+ M+Mk						A1, C4 , D1
9	B+M+ Mk						A2, C5
10	M+ Mk						A3
11	Mk+E1				E1		A5
12	Mk3+ E					Mk3	B3, B4, (B5)

Les résultats sont globalement semblables à ce qui était attendu. Parmi les signatures attendues (partie haute du tableau 4.3, lignes 1 à 7), on trouve les gènes surexprimés uniquement dans un type de leucémie donné (gènes spécifiques des T, B, E, Mk, lignes 1, 2, 4, 5), les gènes spécifiquement surexprimés dans les leucémies lymphoïdes (ligne 3) et les leucémies non-lymphoïdes (lignes 6 et 7). Mais il existe aussi des patrons d'expressions qui mettent en évidence des similitudes inattendues entre certains types de leucémies (partie basse du tableau, lignes 8 à 12). Il est nécessaire de porter attention à ces signatures particulières. Donc, en plus de s'intéresser aux gènes spécifiques de chaque type de leucémies et de différencier les gènes lymphoïdes versus non-lymphoïdes, il est intéressant d'étudier les similarités entre les leucémies Mk, M et B (lignes 9 et 10), mais aussi la similarité de E1 avec les 3 leucémies Mk (ligne 11) et celle de Mk3 avec les 3 leucémies E (ligne 12), et aux gènes uniquement sous-exprimés dans les leucémies E (ligne 8).

4.5. Résultats partie III: gènes spécifiques de chaque type de leucémies

4.5.1 Méthodologie

Les analyses SOM précédentes ont déjà permis d'identifier les groupes de gènes spécifiques de chaque type de leucémies. Cependant une autre méthode plus exacte et largement utilisée existe et elle est nommée SAM (*Significance Analysis of Microarrays*). Cette méthode présente l'avantage de prendre en compte l'ensemble des données et non seulement les 3000 *probesets* sélectionnés. De plus, elle permet le calcul de valeurs statistiques qui permettent une estimation de la valeur de confiance pouvant être accordée aux taux d'expression différentielle. SAM a été développée par le groupe de Tibshirani (Tusher et al. 2001).

La valeur issue de SAM (valeur *d* ou *SAM d-score*) est représentative de la différence d'expression entre les deux conditions testées et prend également en compte les valeurs d'écart type entre les réplifications. Ainsi, plus la valeur *d* est grande et plus il y a une différence élevée entre les 2 conditions et moins il y a de variation entre les réplifications. La valeur *p* issue de SAM teste l'hypothèse nulle que la différence d'expression d'un *probeset* entre les 2 conditions est seulement due au hasard. Plus la valeur *p* est petite et plus l'hypothèse nulle peut être rejetée. A partir de SAM, une autre valeur est calculée et correspond à une estimation du taux de faux positifs pour un *probeset* donné et nommé FDR (*False Discovery Rate*). Ainsi une valeur de FDR de 0,10 signifie une probabilité de 10% que la différence d'expression soit un faux positif. Cette valeur est calculée par permutations des données.

Une analyse SAM a été effectuée grâce au logiciel Flexarray pour obtenir les gènes spécifiques de chaque type de leucémie. Les données issues de SAM (*SAM d-score*) ont été rangées par ordre décroissant pour obtenir une liste ordonnée des gènes les plus spécifiques. Seuls les gènes dont la valeur *p* est inférieure à 0,01 et le FDR inférieur à 0,20 ont été sélectionnés.

L'étape suivante de l'analyse a été d'évaluer la fonction des gènes spécifiques des leucémies. Deux outils ont principalement été utilisés: Netaff (www.affymetrix.com) et le site de NCBI (www.ncbi.nlm.nih.gov). NetAFF fonctionne en faisant le lien entre le

numéro de *probeset* Affymetrix et des entrées dans différentes banques de données. Un des liens est par exemple la banque de données Unigene du site NCBI. Unigene permet alors de connaître le patron d'expression dans différents tissus, la similarité avec des protéines de plusieurs espèces, les articles publiés. Netaff fait aussi le lien entre le gène et les annotations du projet GO (Gene Ontology, <http://www.geneontology.org>). Ainsi le gène est classifié selon sa localisation cellulaire, sa fonction moléculaire et le processus physiologique dans lequel il est impliqué.

Parmi les termes fonctionnels associés aux gènes, certains vont être plus en rapport avec la leucémie ou l'hématopoïèse. Il s'agit de ceux qui peuvent indiquer:

- 1) que le gène code pour une protéine sécrétée (termes fonctionnels: compartiment extracellulaire, cytokine, etc...)
- 2) que le gène code pour un récepteur membranaire (termes fonctionnels: récepteur, localisation membranaire et/ou extracellulaire)
- 3) des molécules de transmission de signaux (notamment les voies de signalisation Ras, Rho, Rab, compartiment cytosolique, etc...)
- 4) des régulateurs du cycle cellulaire et des facteurs de transcription (localisation nucléaire, domaines en doigt de zinc, etc...)

En fait, tout gène dont la fonction n'a pas encore été reliée ni à la leucémie ni au type cellulaire en question est d'intérêt. Une partie des gènes spécifiques de chaque type de leucémie va correspondre à des oncogènes puisque nos échantillons sont leucémiques et une autre partie va être reliée à l'hématopoïèse et à la différenciation cellulaire.

Des essais de classification par ontologie de façon automatisée ont été tentés mais les résultats n'ont pas été concluants. Il s'agit de découvrir les principales classes fonctionnelles présentes dans les données issues d'une l'analyse. Par exemple, le logiciel analyse la répartition fonctionnelle de 500 gènes spécifiques des leucémies T. Les logiciels essayés ont été Genmapp, DAVID (version web de EASE), outil d'ontologie de NettAff. Les résultats ne sont pas convaincants parce que les classes fonctionnelles sont trop larges (par exemple, gènes impliqués dans le métabolisme, impliqués dans la liaison à l'ADN). De plus, beaucoup de gènes codent pour des protéines ayant des fonctions multiples et se retrouvent donc dans plusieurs classes à la fois. Inversement certains gènes n'ont pas encore une fonction connue. Afin de pouvoir être utile à notre étude, les logiciels ont besoin d'être améliorés avec la possibilité d'entrer des mots clés pour

orienter les recherches: par exemple, indiquer avec quel transcriptome on travaille (système hématopoïétique) et/ou indiquer vers quoi est orienté le design expérimental (cancer, etc,...).

Les sections suivantes vont présenter les résultats des analyses pour les leucémies non-lymphoïdes et lymphoïdes. Les leucémies non-lymphoïdes vont être commentées de façon un peu plus approfondie que les leucémies lymphoïdes: d'une part, les leucémies mégacaryoblastiques sont encore très peu caractérisées et il est également nécessaire d'identifier les gènes distinguant les leucémies érythroïdes et mégacaryoblastiques. D'autre part, la caractérisation des leucémies lymphoïdes par l'analyse des micropuces a été le sujet de maîtrise de Cyndia Charfi (étudiante dans le laboratoire d'Eric Rassart) qui poursuit actuellement sur le même sujet au doctorat.

4.5.2. Analyse ciblée sur les leucémies non-lymphoïdes

4.5.2.1 Rappel des caractéristiques des leucémies non-lymphoïdes

Les leucémies mégacaryoblastiques induites par le rétrovirus Graffi ont 2 phénotypes, CD41⁺Kit⁺ et CD41⁺Kit⁻. Dans une même leucémie, les 2 phénotypes co-existent fréquemment. Mk1 représente le phénotype CD41⁺Kit⁺ et Mk2 et Mk3 représentent le phénotype CD41⁺Kit⁻. Les leucémies érythroïdes sont Ter119⁺CD71⁺. Le niveau d'expression de Ter119 peut varier et les leucémies érythroïdes sont parfois Ter119^{low}CD71⁺. Ce phénotype est certainement plus immature car l'expression de Ter119 augmente avec la différenciation. Cependant, nous n'avons trié que les leucémies Ter119⁺CD71⁺ du fait que seul l'anticorps anti-Ter119 était disponible pour trier les leucémies érythroïdes.

Les leucémies érythroïdes et mégacaryoblastiques induites par le rétrovirus Graffi sont particulièrement intéressantes à étudier parce que les leucémies homologues chez l'humain, AML-6 et AML-7, sont associées à de très faibles chances de guérison et sont moins bien connues que les leucémies de type T ou B. De plus, au-delà de la notion d'oncogenèse, il est utile de mieux connaître les gènes spécifiques de ces 2 lignées. Elles sont très proches et ont notamment beaucoup de facteurs de transcription en commun et il est essentiel de mieux connaître ce qui les différencie (chapitre I, section 1.1.1.6.2.2).

4.5.2.2. Vérification des phénotypes

Le tableau 4.4 ci-dessous permet de vérifier le phénotype des leucémies érythroïdes et mégacaryoblastiques.

Tableau 4.4: vérification des phénotypes E et Mk													
E													
	T1	T2	T3	B1	B2	B3	M	E1	E2	E3	Mk1	Mk2	Mk3
Gypa	-2,40	-1,94	-2,42	-3,07	-2,78	-2,59	-1,61	5,68	5,46	4,80	-1,95	-0,71	3,53
Gypa	-2,28	-1,86	-2,08	-2,13	-2,27	-1,99	-1,08	5,02	4,37	4,19	-1,61	-1,02	2,74
Tfrc	-0,23	0,86	0,30	-0,01	0,66	0,06	-1,71	1,52	1,63	1,14	-1,27	-2,42	-0,52
Tfrc	-0,06	0,87	0,78	-0,27	0,86	0,13	-1,85	1,71	1,92	1,36	-2,08	-2,28	-1,09
Epor	-1,60	-0,61	-0,95	-1,62	-1,54	-1,38	-1,55	3,14	3,12	2,65	0,02	-0,92	1,25
MK													
Itga2b	-2,71	-1,62	-2,56	-2,88	-2,58	-2,74	-1,62	3,78	0,81	1,13	3,63	3,58	3,78
Kit	-2,19	-2,13	-2,17	-2,20	-2,34	1,18	2,84	-1,68	-1,60	-2,15	3,76	4,64	4,04
Kit	-1,52	-1,46	-1,31	-1,56	-1,46	0,03	2,14	-1,04	-1,13	-1,42	2,71	3,75	2,26
Mpl	-1,89	-1,61	-2,11	-1,98	-2,09	-1,96	-1,33	1,29	-1,00	-1,74	4,85	4,40	5,17
Mpl	-0,90	-0,85	-0,82	-0,53	-0,72	-0,67	-0,82	-0,37	-0,60	-0,74	2,55	1,71	2,76
Cxcl4	-3,12	-0,95	-3,62	-3,69	-3,99	-3,37	1,15	4,78	-1,70	1,93	3,90	4,00	4,68

La première partie du tableau montre les gènes spécifiques de la lignée érythroïde, soit Gypa, Tfrc et Epor. La deuxième partie du tableau montre les gènes spécifiques de la lignée mégacaryocytaire, soit Itga2b, Kit, Mpl et Cxcl4. Certains gènes sont représentés par 2 lignes qui correspondent à des *probesets* redondants pour chaque gène (les 2 *probesets* sont placés à 2 régions différentes du même gène, par exemple, dans le 3'UTR et dans les exons, permettant aussi de différencier les transcrits alternatifs).

Le gène GypA code pour la protéine glycophorine A, Ter119 étant un épitope de cette protéine. Tfrc est le gène du récepteur de la transferrine soit CD71. Ainsi ces résultats confirment bien le phénotype Ter119⁺CD71⁺ des leucémies érythroïdes. Epor code pour le récepteur à l'érythropoïétine et est donc le gène spécifique de la lignée érythroïde (Richmond et al. 2005). On peut remarquer que Mk3 exprime aussi Gypa et Epor confirmant les caractéristiques érythroïdes de Mk3, ce qui avait été remarqué par l'étude du profil des données (section 4.3.3.3). Cependant l'expression de Ter119 n'a pas été détectée en cytométrie de flux sur Mk3, l'expression du gène n'étant peut-être pas corrélée avec la traduction de la protéine.

Tfrc (CD71) est aussi un peu exprimé par les leucémies T et B. Ceci est en corrélation avec toutes les analyses de cytométrie en flux réalisées sur les leucémies induites avec le rétrovirus Graffi et c'est pour cette raison que seul l'anticorps spécifique à Ter119 permet de trier les leucémies érythroïdes.

Itga2b code pour la glycoprotéine IIb qui correspond à CD41, et qui forme l'intégrine $\alpha_{IIb}\beta_3$ avec Itga3b (CD61). D'après les résultats des micropuces, Itga2b serait fortement exprimé dans les 3 leucémies érythroïdes et les 3 leucémies mégakaryoblastiques. Cependant, d'après les résultats de la cytométrie en flux, seulement Mk2 et Mk3 seraient CD41⁺ (tableau 4.1). Ainsi, dans ce cas également, on peut penser qu'il y a un manque de corrélation entre l'expression du gène et la traduction de la protéine ou bien la conformation non active de l'intégrine $\alpha_{IIb}\beta_3$ ne permet pas la détection par l'anticorps anti-CD41. Ce phénomène a déjà été observé pour ce gène lors de la caractérisation des leucémies induites par le rétrovirus Graffi (chapitre II) et on peut ajouter que l'anticorps CD41 ne reconnaît que la forme mature et active de l'intégrine $\alpha_{IIb}\beta_3$.

L'expression du gène Kit est, d'après les résultats des micropuces, en accord avec les analyses de cytométrie en flux et exprimé fortement dans les leucémies mégakaryoblastiques et la leucémie myéloïde. Mpl et Cxcl sont 2 gènes supplémentaires spécifiques de la lignée mégacaryocytaire qui avaient déjà été testés lors de la caractérisation des leucémies induites par le rétrovirus Graffi et les résultats des micropuces sont en accord avec ces résultats antérieurs (chapitre II). Mpl code pour le récepteur de la thrombopoïétine et Cxcl code pour la protéine *Platelet factor 4*. Mpl est fortement surexprimé dans les 3 leucémies Mk et à moindre degré dans E1. Cxcl4 est surexprimé dans les 3 leucémies Mk, la leucémie myéloïde, E1 et à moindre degré E3. L'expression de Mpl et Cxcl4 par E1 montre les caractéristiques mégakaryoblastiques de cet échantillon (section 4.3.3.3).

Les analyses des micropuces permettent également d'identifier d'autres marqueurs spécifiques (CD, Cluster of Differentiation) surexprimés par les leucémies érythroïdes ou mégakaryoblastiques induites par le rétrovirus Graffi (appendice A). Il est intéressant de connaître ces marqueurs supplémentaires car ils peuvent être utilisés en analyse de cytométrie en flux ou pour trier les cellules.

Cd44 et Cd47 sont spécifiques des leucémies érythroïdes et mégacaryoblastiques induites par le rétrovirus Graffi (appendice A). Cd44 est connu pour être à la surface des érythrocytes mais aussi des leucocytes (Telen et al. 2005). De façon intéressante, il est aussi un marqueur spécifique des cellules souches leucémiques des leucémies AML (Jin et al. 2006). CD47 se trouve à la surface des érythrocytes (Ishikawa-Sekigami, 2006) et des plaquettes (Lagadec et al. 2003).

Les marqueurs spécifiques supplémentaires pour les leucémies mégacaryoblastiques sont Cd151, Cd14, Cd68, Cd83, Cd96, Cd9 (appendice A). Cd151 code pour une protéine de la famille des tétraspanines et son expression est connue dans les mégacaryocytes et les plaquettes. Une étude montre une interaction entre CD151 et l'intégrine $\alpha_{\text{IIb}}\beta_3$ (CD41/CD61) et CD151 aurait un effet sur l'acquisition de la conformation active de l'intégrine (Lau et al. 2004). Cd9 qui code aussi pour une tétraspanine est aussi connue pour interagir avec l'intégrine $\alpha_{\text{IIb}}\beta_3$ dans les plaquettes (Longhurst et al. 1999). CD14 est classiquement un antigène myéloïde, présent à la surface des macrophages et également des cellules dendritiques. CD14 interagit avec TLR4 (Toll-like receptor 4), ce complexe interagissant avec les molécules de LPS bactériens. Les 3 autres marqueurs CD68, CD83 et CD96 ne sont pas connus pour être spécifiques des mégacaryocytes et plaquettes mais ont comme point commun celui d'être présent à la surface de cellules néoplasiques. Cd68 est fortement exprimé par les monocytes/macrophages et il est aussi surexprimé en cas de néoplasie (Strobl et al. 1995). Cd83 est connu pour être un marqueur des cellules dendritiques et il a également été identifié dans certains types de cancers chez l'humain (Baleeiro et al. 2007). Cd96 a été détecté sur les cellules NK et T mais très récemment il a été rapporté spécifique des cellules souches leucémiques des leucémies AML chez l'humain (Hosen et al. 2007).

Les marqueurs supplémentaires spécifiques des leucémies érythroïdes sont Cd24, Cd59a, Cd7 et à moindre degré Cd48 (appendice A). Cd59a et Cd24 sont connus pour être exprimés à la surface des érythrocytes murins (Meri et al. 1990, Schabath et al. 2005). Cd7 est exprimé sur les progéniteurs myéloïdes et lymphoïdes et en cas de leucémies (Sempowski et al. 1999). Enfin, CD48 est également restreint aux progéniteurs myélo-érythroïdes et aux cellules B (Kiel et al. 2005).

Ainsi, certains de ces marqueurs sont connus pour être à la surface des érythrocytes ou des plaquettes alors que d'autres ne le sont pas. Ces derniers attestent peut-être de l'état cancéreux des cellules, ou de leur stade immature. Il est important de vérifier ces expressions sur les leucémies provenant d'autres modèles et notamment chez l'humain surtout pour les leucémies mégacaryoblastiques qui sont moins connues.

4.5.2.3. Exemples de gènes connus pour les leucémies mégacaryoblastiques et les leucémies érythroïdes

Bien qu'à l'état cancéreux, les cellules leucémiques conservent beaucoup de caractéristiques des cellules dont elles dérivent. Dans cette section, les expressions de gènes connus pour être impliqués soit dans la lignée érythroïde soit mégacaryocytaire vont être analysées. Les gènes choisis sont les plus typiques de ces lignées. Cela permet de vérifier que les caractéristiques essentielles des cellules sont conservées et qu'il s'agit donc de bons modèles pour étudier ces types de leucémies.

4.5.2.3.1. Facteurs de transcription

Tableau 4.5: facteurs de transcription spécifiques des E et Mk													
E et MK													
	T1	T2	T3	B1	B2	B3	M	E1	E2	E3	Mk1	Mk2	Mk3
Gata1	-1,87	-1,92	-2,07	-2,07	-2,31	-2,65	-1,89	4,49	4,15	4,48	0,58	-1,44	2,52
Gata2	-1,87	-1,61	-1,91	-1,67	-1,82	-1,94	-1,05	-0,57	2,19	0,62	3,42	3,44	2,77
Gata2	-2,29	-2,01	-2,29	-2,13	-2,39	-1,93	-1,63	0,15	2,41	1,08	3,63	4,02	3,39
Zfpml(Fog1)	-1,18	-1,74	-1,39	0,11	-0,59	-0,23	-3,95	2,00	2,01	1,95	1,27	0,79	0,95
Zfpml(Fog1)	-1,11	-1,93	-0,97	-0,02	-0,80	-0,28	-2,62	1,82	1,94	1,86	1,12	0,38	0,61
Tal1	-3,26	-2,50	-3,10	-3,25	-2,55	-2,52	-2,49	4,07	3,75	3,46	2,56	2,94	2,87
Lmo2	-5,10	-3,79	-4,63	0,74	-1,33	-0,87	0,70	2,82	3,01	2,90	1,30	2,47	1,76
Nfe2	-4,22	-3,04	-3,44	-3,81	-3,00	-3,16	1,33	3,68	3,45	3,28	2,47	3,66	2,80
Fli1	-0,23	-0,37	-0,48	0,70	0,58	0,47	0,70	2,56	-4,82	-4,93	1,84	2,36	1,62
Fli1	-0,62	0,53	-2,01	0,36	0,00	-0,07	-0,46	2,39	-2,92	-1,66	1,32	1,88	1,26
E													
Klf1	-2,91	-2,34	-2,37	-2,71	-2,75	-2,84	-1,81	4,81	5,45	5,33	1,72	-1,65	2,07
Trim10	-1,96	-1,43	-2,12	-2,15	-2,08	-1,97	-1,52	4,99	4,08	3,79	-0,97	-0,81	2,14
MK													
Runx1	0,55	-1,81	-1,04	0,16	0,21	0,46	-0,45	0,92	-0,75	-0,06	0,71	0,79	0,30
Runx1	0,60	-1,88	-0,68	-0,01	0,15	0,41	-0,64	0,83	-0,61	0,25	0,70	0,82	0,07

Le tableau ci-dessus (tableau 4.5) montre les résultats des micropuces pour certains facteurs de transcription impliqués dans l'érythropoïèse et/ou la mégacaryopoïèse (Chapitre I, section 1.1.1.6.2.2).

L'expression de ces facteurs de transcription dans les leucémies induites par le rétrovirus Graffi est en accord avec leurs rôles joués dans l'érythropoïèse et la mégacaryopoïèse (Chapitre I, section 1.1.1.6.2.2). Gata2, Fog1, Tal1 (Scl), Lmo2 et Nfe2 sont exprimés fortement dans les leucémies érythroïdes et mégacaryoblastiques. Gata1 semble plus érythroïde, Fli-1 plus mégacaryoblastique. Klf (Eklf) et Trim10 (Herf1), connus pour agir sur les stades tardifs de l'érythropoïèse (chapitre I), sont bien exprimés le plus fortement dans les 3 leucémies érythroïdes. L'expression de ces 2 gènes par Mk3 montre encore une fois les caractéristiques érythroïdes de cet échantillon. Runx1 (AML1) est plus spécifique des leucémies mégacaryoblastiques et son expression dans E1 fait probablement partie des caractéristiques mégacaryoblastiques de E1.

4.5.2.3.2. L' intégrine plaquettaire $\alpha\text{IIb}\beta_3$

Les protéines codées par les gènes Itga2b (CD41) et Itgb3 (CD61) forment un dimère appelé intégrine $\alpha\text{IIb}\beta_3$ correspondant au récepteur du fibrinogène et responsable de l'agrégation plaquettaire. Des fibres de fibrinogène vont se lier par leurs différentes extrémités aux intégrines $\alpha\text{IIb}\beta_3$ de plusieurs plaquettes. Cela va former un réseau plaquettaire (thrombose). Il va s'ensuivre le processus de coagulation ainsi que l'attraction des leucocytes sur le site. L'agrégation plaquettaire est précédée par l'activation plaquettaire et la sécrétion des granules. Ceci est un phénomène hautement régulé et donc de nombreux signaux arrivent et partent de l'intégrine $\alpha\text{IIb}\beta_3$. Certaines protéines sont connues pour interagir avec cette intégrine, pour réguler, ou être régulées par elle.

La figure 4.8 A montre un schéma de l'intégrine $\alpha\text{IIb}\beta_3$ avec ses principaux partenaires et le tableau 4.6 récapitule les résultats des micropuces pour les différents gènes correspondants.

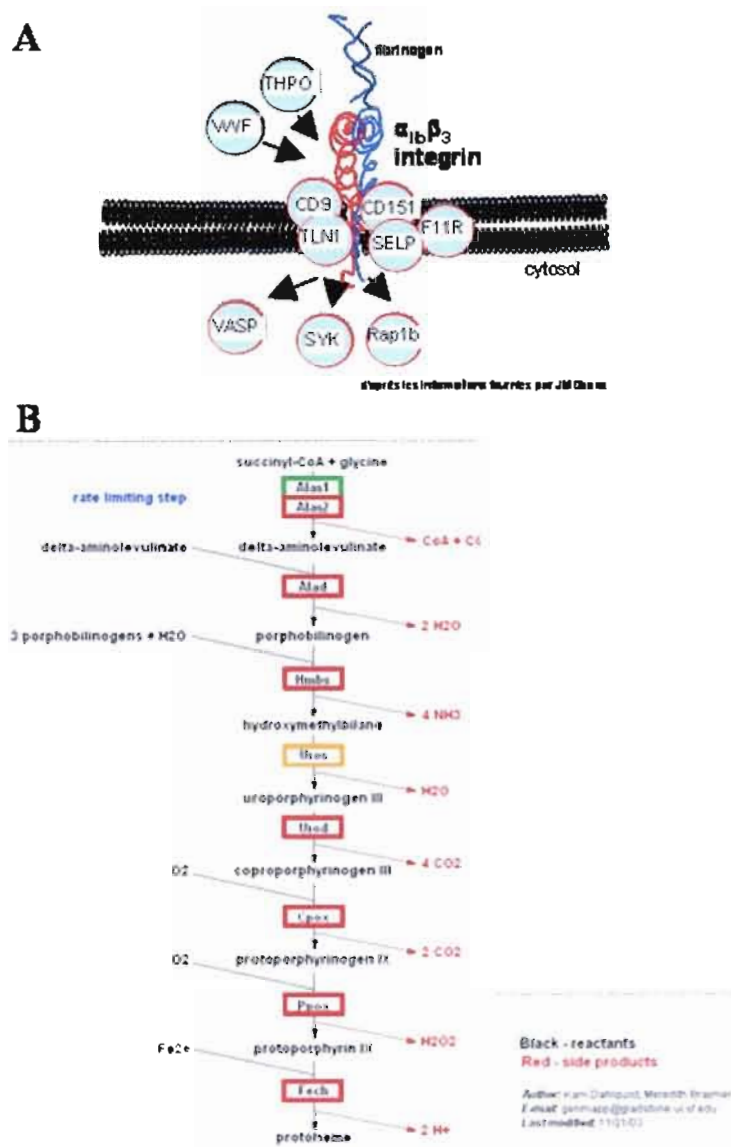


Figure 4.8: Schéma du complexe de l'intégrine $\alpha_{IIb}\beta_3$ et de la biosynthèse du hème. (A) Représentation de l'intégrine $\alpha_{IIb}\beta_3$ avec ses partenaires d'interaction connus. Les cercles rouges indiquent une surexpression des gènes correspondants dans les leucémies mégacaryoblastiques induites par le rétrovirus Graffi. (B) Biosynthèse du hème, schéma tiré de Keggpathway. Cadre rouge: forte surexpression dans les leucémies érythroïdes induites par le rétrovirus Graffi; cadre orange: faible surexpression; cadre vert: sous-expression.

	T1	T2	T3	B1	B2	B3	M	E1	E2	E3	Mk	MK	MK3	
Itga2b	-2,7	-1,6	-2,6	-2,9	-2,6	-2,7	-1,6	3,8	0,8	1,1	3,6	3,6	3,8	integrin alpha 2b
Itgb3	-1,0	-0,7	-0,7	-0,8	-0,7	-0,9	-0,6	1,6	-0,7	-0,7	-0,6	2,9	3,1	integrin beta 3
Cd9	1,1	-2,5	2,5	-3,5	-3,6	-3,6	0,6	1,9	-1,8	-2,1	3,0	4,4	3,5	CD9 antigen
Cd151	-1,2	-1,1	-0,7	-0,7	-0,9	-0,6	-0,3	1,4	-0,2	-0,3	1,8	0,8	2,1	CD151 antigen
Selp	-1,0	-1,2	-1,9	-1,6	-1,5	-1,7	-1,1	-0,9	-1,7	-1,6	4,3	5,7	4,1	selectin, platelet
Tln1	-0,7	-0,4	-1,4	-0,3	-0,8	-1,0	0,9	2,3	-1,5	-0,5	0,8	1,4	1,3	Talin 1
F11r	-0,9	-0,8	-0,9	-0,8	-0,6	-1,0	-0,6	0,2	-0,6	-0,6	2,0	3,1	1,5	F11 receptor
vasp	-0,8	-1,3	0,1	0,3	-0,3	0,3	0,7	0,0	-0,6	-1,0	0,8	1,2	0,7	vasodilator-stimulated phosphoprotein
Syk	-2,4	-2,8	-2,8	2,9	2,5	2,8	1,4	-1,5	-2,0	-3,1	1,5	1,6	1,8	spleen tyrosine kinase
Rap1b	-0,7	-0,8	-1,1	-0,6	-1,2	-1,0	-0,2	2,1	-0,3	0,9	0,7	0,9	1,4	RAS related protein 1b
F5	-0,7	-0,8	-0,4	-0,7	-0,7	-1,0	2,0	0,3	-0,6	-0,6	-0,4	2,3	1,2	coagulation factor V
F10	-1,0	-1,0	-1,1	-1,5	-1,2	-1,7	5,0	-1,1	-1,2	-1,3	-0,3	4,1	2,5	coagulation factor X
F13a1	1,0	-2,1	-2,2	-0,7	0,1	-1,7	5,7	-0,8	-1,8	-3,2	-0,6	3,6	2,8	coagulation factor XIII, A1 subunit
F2r	-1,2	-0,3	-3,2	-4,5	-3,6	-3,3	-2,2	1,6	2,1	2,4	3,7	4,5	4,0	coagulation factor II receptor

Les interactions de CD151 et CD9 avec l'intégrine $\alpha_{\text{Iib}}\beta_3$ ont été rapportées précédemment (Vérification des phénotypes, section 4.5.2.2).

L'activation du récepteur codé par F11r stimule la dégranulation et l'agrégation plaquettaire. F11R est connu pour interagir physiquement avec l'intégrine $\alpha_{IIb}\beta_3$ et également CD9 (Sobocka et al. 2004).

SYK, une tyrosine kinase non-récepteur, et RAP1b, une GTPase, sont activées par l'intégrine $\alpha_{IIb}\beta_3$ et envoient à leur tour des signaux à l'intérieur de la cellule (Gao et al. 1997, Bertoni et al. 2002).

Une trop grande ampleur du phénomène d'agrégation aurait des effets néfastes sur l'organisme. Ainsi dès l'activation, des signaux de régulation négative sont aussi envoyés à l'intégrine $\alpha_{\text{IIb}}\beta_3$. VASP qui interagit avec le cytosquelette est un exemple d'un régulateur négatif de l'intégrine $\alpha_{\text{IIb}}\beta_3$ (Hauser et al. 1999).

L'étape suivante après l'agrégation plaquettaire est la coagulation, processus auquel participent également les plaquettes. Le résultat des micropuces montre, dans le bas du tableau 4.6, la surexpression de 3 gènes codant pour des facteurs de coagulation (F5, F10 et F13) par les leucémies mégacaryoblastiques. La cascade de coagulation permet la production de thrombine active et l'étape finale de la cascade est la production de molécules de fibrines à partir du fibrinogène. Les molécules de fibrines vont former un caillot sanguin solide. Les leucémies mégacaryoblastiques surexpriment également le récepteur à la thrombine (F2r) (tableau 4.6). La thrombine est un stimulateur puissant de l'activation plaquettaire et le récepteur à la thrombine régule notamment de façon positive l'intégrine $\alpha_{\text{IIb}}\beta_3$ (Voss et al. 2007). F2r est également fortement surexprimé dans les leucémies érythroïdes induites par le rétrovirus Graffi. Aucune information bibliographique n'a pu confirmer l'expression de ce récepteur dans la lignée érythroïde non leucémique.

4.5.2.3.3. La synthèse de l'hème par les érythrocytes

Les globules rouges ont pour fonction de transporter l'oxygène de l'organisme. Ainsi, un globule rouge mature énucléé est compris à environ 85% d'hémoglobine. Chaque molécule d'hémoglobine est constituée de 2 chaînes de globine α et de 2 chaînes de globine β . Chaque chaîne de globine comporte un hème contenant une molécule de fer qui lie l'oxygène. La biosynthèse des chaînes de globine et du hème est donc essentielle au bon fonctionnement des globules rouges et elle s'effectue lors de la différenciation, au stade érythroblaste avant que le globule rouge soit énucléé. Dans l'organisme, 85% de la biosynthèse du hème est réalisé par les érythroblastes. La figure 4.8 B montre le schéma de la biosynthèse du hème en mettant en évidence les enzymes responsables. Tous les gènes correspondant à ces enzymes excepté *Alas1* sont surexprimés spécifiquement par les leucémies érythroïdes induites par le rétrovirus Graffi (tableau 4.7).

Tableau 4.7: Synthèse du hème par les érythrocytes														nom complet
titre	T1	T2	T3	B1	B2	B3	M	E1	E2	E3	Mk1	Mk2	Mk3	
Alas1	-0,1	-0,3	-0,5	-0,1	-0,3	-0,7	2,9	-0,4	-0,2	-0,3	-0,4	0,6	-0,4	aminolevulinic acid synthase 1
Alas2	-3,4	0,3	-0,9	-2,5	-3,8	-3,2	-1,3	4,9	4,9	4,8	-2,8	-0,5	3,5	aminolevulinic acid synthase 2, erythroid
Alad	-1,2	-0,4	0,0	-1,3	-1,6	-0,7	-1,6	2,2	2,2	2,0	0,4	-0,9	1,2	aminolevulinate, delta-, dehydratase
Hmbs	-1,2	-1,0	-0,4	-1,5	-1,1	-1,1	-1,1	2,6	3,4	2,7	-0,8	-0,9	0,5	hydroxymethylbilane synthase
Uros	-0,6	-0,3	-0,2	-0,3	0,1	0,2	-0,9	0,7	1,5	0,2	0,0	-0,5	0,0	uroporphyrinogen III synthase
Urod	0,0	-0,4	-0,3	-1,2	-0,9	-1,0	-0,3	1,8	1,6	1,4	-0,6	-0,5	0,3	uroporphyrinogen decarboxylase
Cpox	-1,6	-1,2	-0,2	-0,5	-0,8	-1,0	-1,5	2,7	3,1	3,2	-1,0	-0,7	-0,4	coproporphyrinogen oxidase
Ppox	-0,5	-0,1	-0,3	0,1	-0,2	-0,4	0,0	0,7	1,4	0,6	-0,6	-0,4	-0,4	protoporphyrinogen oxidase
Fech	-0,7	-1,0	-0,3	-1,9	-1,9	-2,0	-0,4	2,5	3,1	2,2	-0,7	0,0	1,1	ferrochelatase
Hbb-b1,2	1,4	0,8	-0,7	-3,7	-4,4	0,5	2,4	2,4	2,3	0,3	1,1	2,1		hemoglobin, beta adult major chain
Hba-a1	1,8	0,7	-1,1	-4,1	-5,3	-0,1	3,0	3,1	3,0	0,1	1,3	2,6		hemoglobin alpha, adult chain I

Les 2 chaînes α et β de la globine sont également fortement exprimées dans les 3 leucémies érythroïdes (Hbb et Hba). On peut néanmoins observer une forte expression dans T2, T3, Mk2 et Mk3. On peut envisager un caractère biphénotypique des leucémies mégacaryoblastiques. Cependant, dans le cas des leucémies T, il est probable qu'il s'agisse d'une contamination par des globules rouges résiduels. En effet, bien qu'énucléé, le globule rouge conserve certaines espèces d'ARN messagers, les molécules d'ARN de globines étant de loin les plus représentées.

En dernier lieu, en plus de la synthèse de l'hème et des globines, un grand apport en fer doit avoir lieu par les érythroblastes. Le fer est complexé à la transferrine dans le plasma. Le complexe fer/transferrine pénètre dans les cellules grâce au récepteur à la transferrine, dont le gène (Tfrc, Cd71) est fortement surexprimé dans les leucémies érythroïdes (tableau 4.4).

Ainsi les leucémies érythroïdes et mégacaryoblastiques induites par le rétrovirus Graffi se distinguent bien l'une de l'autre et des autres types de leucémies car elles surexpriment spécifiquement les gènes qui leur sont les plus caractéristiques.

4.5.2.4. Gènes spécifiques des leucémies mégacaryoblastiques

D'après le profil général des leucémies (section 4.3.3.3), il a été observé que E1 a des caractéristiques mégacaryoblastiques et que les leucémies Mk ont aussi beaucoup de gènes communs avec les leucémies M et B. Ainsi, 4 tableaux ont été réalisés pour trouver les gènes les plus différenciellement exprimés et donc les plus spécifiques des leucémies mégacaryoblastiques comprises dans cette étude (méthodologie décrite dans la section 4.5.1). Ces tableaux permettent notamment de sélectionner des gènes candidats dont il

serait très intéressant d'étudier la fonction dans les leucémies mégacaryoblastiques ou plus généralement dans la lignée mégacaryocytaire. Ces tableaux sont présentés et commentés un à un dans les sections suivantes.

4.5.2.4.1 Gènes surexprimés dans MK1 MK2 MK3

Le premier tableau (appendice B) présente les gènes uniquement spécifiques des leucémies Mk1, Mk2 et Mk3. A l'issue de l'analyse SAM, 920 *probesets* ont passé les critères d'une valeur p inférieure à 0,01 et une valeur de FDR inférieure à 0,20. Les *probesets* ont été organisés de façon décroissante selon les valeurs issues de SAM (*SAM d-score*) pour obtenir une liste organisée des *probesets*.

Ce tableau (appendice B) présente les résultats des 50 premiers *probesets* les plus spécifiques des leucémies mégacaryoblastiques. Pour des raisons de mise en page, seuls les 50 premiers *probesets* de ce tableau sont montrés mais cela est très réducteur puisque les valeurs statistiques de SAM indiquent que 920 *probesets* ont potentiellement une expression qui varie de façon significative. En pratique, l'analyse s'est effectuée sur approximativement les 200 premiers *probesets*.

Le tableau (appendice B) est organisé en 2 parties. Sur la première partie, la première colonne indique l'ordre des *probesets*, et la deuxième colonne le nom des *probesets* (colonne 'F1'). Les 4 colonnes qui suivent correspondent aux valeurs issues de SAM et les 13 colonnes suivantes indiquent les résultats des ratios d'expression (en rouge: surexpression significative du gène par rapport à la moyenne des échantillons; en vert: sous expression). Il s'ensuit le nom officiel des gènes correspondant aux *probesets* ('*gene symbol*'). Les gènes colorés en jaune sont représentés par plusieurs *probesets* (*probesets* redondants) dans les résultats de SAM et l'écriture en violet montre que ces gènes ont été choisis pour être validés en RT-PCR. La dernière colonne marque les gènes qui sont déjà connus dans la lignée mégacaryocytaire (mégacaryoblastes, mégacaryocytes, plaquettes). Sur la deuxième partie du tableau (appendice B suite), le nom des gènes est répété, suivi du nom complet des gènes. Les 3 colonnes suivantes indiquent l'ordre des *probesets*, le nombre de *probesets* redondants pour chaque gène, et l'ordre des gènes compte tenu des *probesets* redondants. La redondance des *probesets* pour un gène est une information très importante car elle reflète l'expression différentielle pour le gène en question. La colonne suivante présente les homologues entre les protéines codées par les gènes d'intérêt et les

protéines humaines (% homologie H.s). Un pourcentage supérieur à 50% fait penser que la protéine murine et humaine sont les mêmes. Cette information est essentielle pour la poursuite des travaux puisque le but est d'appliquer les résultats aux leucémies humaines. Les colonnes suivantes présentent les résultats du signal de MAS5.0 (A; 'absent'; P: 'présent'; M: 'medium').

Les 50 premiers *probesets* de ce tableau correspondent à 37 gènes. Les ratios d'expression sont élevés et les résultats de MAS5.0 montrent en général une 'présence' dans les leucémies mégacaryoblastiques et une 'absence' dans les autres échantillons. On peut donc émettre l'hypothèse que les variations d'expression sont très grandes entre les échantillons mégacaryoblastiques et les autres types de leucémies et que dans de nombreux cas, le gène peut ne pas être exprimé du tout par les autres échantillons.

Sur les 50 premiers *probesets*, 8 gènes seulement sont connus pour être exprimés dans la lignée mégacaryocytaire (Fcer1a, Sdpr, Meis1, Mpl, Slc35d3, Selp, Kit, Angpt). Ce nombre peu élevé de gène reflète le fait que les leucémies mégacaryoblastiques sont peu connues et démontre l'importance de les étudier.

Fcer1a (rang 2) fait partie de la famille des récepteurs de reconnaissance des antigènes comme le TCR ou le BCR (récepteur de l'immunoglobuline E). Il reconnaît les bactéries et envoie des signaux d'activation à l'intérieur de la cellule. Il est exprimé à la surface des cellules myéloïdes (éosinophiles, basophiles, monocytes), des cellules dendritiques, de Langerhans et mastocytaires. Une étude a montré également la présence de ce récepteur dans le cytoplasme des mégacaryocytes humains et à la surface des plaquettes humaines (Hasegawa et al. 1999).

Selp (rang 29) code pour la protéine sélectine P dont les rôles dans l'activation plaquettaire et son interaction avec l'intégrine $\alpha_{\text{IIb}}\beta_3$ ont été discutés précédemment (section 4.5.2.3.2). La protéine codée par Sdpr (rangs 3, 5, 46) a été isolée à partir de plaquettes humaines (Burgener et al. 1990) mais peu de choses sont connues sur sa fonction et il serait intéressant de l'approfondir. Slc35d3 (rang 13) code pour un récepteur orphelin et un *knock-out* de ce gène a permis de mettre en évidence son rôle dans la régulation des granules denses des plaquettes (Chintala et al. 2007).

Parmi ces 8 gènes, 3 correspondent à des oncogènes déjà répertoriés. Il s'agit de Meis1 (rang 8), Mpl (rangs 10, 68), et Kit (rang 39).

Meis1 a été découvert à l'origine à proximité d'un site commun d'insertion rétrovirale dans des souris BXH-2 (Moskow et al. 1995). La protéine MEIS1 est connue pour interagir et former un hétérodimère avec PBX1 (Chang et al. 1997). Le gène Pbx1 est également un oncogène bien répertorié et le résultat des micropuces montre qu'il est également surexprimé dans les leucémies mégacaryoblastiques induites par le rétrovirus Graffi (non montré, rangs 51, 57, 173, 174, 548). Meis1 et Pbx1 ont notamment été caractérisés spécifiques des leucémies mégacaryoblastiques murines induites par le rétrovirus murin Cas-Br-E (Hao et al. 2006). Le grand nombre de *probesets* pour ces gènes confirme également le caractère très spécifique de ces 2 oncogènes pour les leucémies mégacaryoblastiques induites par le rétrovirus Graffi. Pbx1 est impliqué dans une translocation chromosomique impliquant le gène E2A (Dedera et al. 1993). MEIS1 et PBX1 régulent ensemble le gène Cxcl4 (Platelet factor 4, voir tableau 4.4) spécifique de la lignée mégacaryocytaire (Okada et al. 2003). Ils sont tous les deux impliqués dans l'établissement de l'hématopoïèse définitive (DiMarino et al. 2001, Azcoitia et al. 2005) et notamment le *knock-out* de Meis1 résulte en une absence de formation de la lignée mégacaryocytaire (Azcoitia et al. 2005). De façon intéressante, des études ont prouvé que PBX1 se fixe au niveau de U5 du rétrovirus Moloney sur la séquence TGATTGAC avec une action positive sur la transcription rétrovirale. Cette séquence est parfaitement conservée sur U5 de GV-1.2 et GV-1.4 (non montré) et on peut donc penser que PBX1 aiderait aussi à la transcription virale du rétrovirus Graffi. Il serait intéressant de voir si Pbx1 et Meis1 sont aussi spécifiques des AMKL chez l'humain. La grande homologie entre les protéines souris et humaines (89% pour Meis1 et 100% pour Pbx1) encourage à penser qu'elles ont chez l'humain et chez la souris une fonction très similaire.

Mpl (myeloproliferative leukemia virus oncogene) n'est autre que le récepteur de la thrombopoïétine, le régulateur principal de la production des plaquettes par les mégacaryocytes (Vainchenker et al. 1995). Il est important pour la différenciation mégacaryocytaire. Il est donc logique de voir apparaître cet oncogène dans les premiers rangs de l'appendice B. Mpl est un membre de la superfamille des récepteurs de cytokines et il a été découvert comme partie intégrante d'un virus transformant de Friend qui induisait préférentiellement des leucémies de type myéloïde (myéloïdes, érythroïdes, mégacaryoblastiques et mastocytaires) (Souyri et al, 1990). Son potentiel oncogénique intrinsèque a aussi été démontré par sa capacité à transformer des fibroblastes de rat (Challier et al, 2000).

Trois oncogènes connus mais non répertoriés dans la lignée ou leucémie mégacaryoblastiques sont Nmyc (rang 11), Evi3 (Zfp521, rang 22) et Fgf3 (rangs 14, 43).

Nmyc est aussi un oncogène partageant une homologie avec c-Myc. Il est capable de transformer des cellules normales en culture en coopération avec Ras (Yancopoulos et al. 1985). Il est un marqueur d'un type de cancer particulier, le neuroblastome (Schwab et al. 1983). Il a été retrouvé dans de nombreux cas de leucémies (Hirvonen et al. 1991) et est la cible d'insertions rétrovirales (van Lohuizen et al. 1989). Son expression n'a pas été montrée dans les tissus hématopoïétiques normaux (Hirvonen et al. 1981). Il n'a pas encore été identifié ni dans les mégacaryocytes ni dans les leucémies mégacaryoblastiques. Ainsi il serait intéressant de vérifier les résultats des micropuces et voir s'il existe une surexpression dans des leucémies AMKL humaines. Sa faible expression dans les tissus sains en fait une cible privilégiée pour une approche thérapeutique.

Evi3 est la cible d'un site commun d'intégration dans des lymphomes B de souris AKXD (Hentges et al. 2005) et la dérégulation de ce gène entraîne la tumorigenèse. L'homologue humain de Evi3, EHZF, est trouvé dans les cellules hématopoïétiques CD34⁺ uniquement (Bond et al. 2004). De plus, son expression a été testée dans plusieurs lignées leucémiques humaines, T, B et myéloïdes et une forte expression a été détectée seulement dans les lignées HEL et K562, qui proviennent de leucémies myéloïdes avec un phénotype érythro-mégacaryocytaire. De nombreux transcrits alternatifs existent pour Evi3 et 2 seulement des 5 *probesets* présents sur la micropuce donnent une spécificité pour les leucémies Mk.

Fgf3 (Fibroblast growth factor 3) est la cible d'un site commun d'intégration rétrovirale du virus MMTV (Dickson et al, 1990) et d'un autre site commun dans des leucémies induites par le rétrovirus Friend (Fis-1, Silver et al, 1986). Sa fonction normale semble être reliée au développement. Son rôle oncogénique est prouvé (Venesio et al, 1992, Kohl et al. 2002) et il est impliqué dans des cancers humains (Hu et al. 2007). Il est relativement peu étudié et il serait intéressant, à la vue des résultats des micropuces, de valider les résultats et de poursuivre son étude fonctionnelle en relation avec les leucémies AMKL. Il fait partie de la grande famille des '*fibroblast growth factors*' qui est connue pour moduler la croissance, la différenciation, la migration ou la survie cellulaire.

Il est également intéressant de noter la surexpression de 2 autres oncogènes même s'ils sont situés au-delà des 50 premiers *probesets*. Il s'agit de Fap et Evi1.

Le gène FAP (rang 151) code pour une protéine qui est exprimée par plus de 90% des fibroblastes tumoraux entourant des carcinomes épithéliaux humains (Garin-Chesa et al. 1990) et son rôle comme inducteur de tumeur a récemment été prouvé (Cheng et al. 2005). Une molécule Val-boroPro a été trouvée capable d'inhiber l'activité enzymatique de FAP ainsi que son effet tumorigène et pourrait peut-être utilisée comme agent thérapeutique (Cheng et al. 2005). Ce gène étant une cible potentielle pour la thérapie, il serait intéressant de le valider sur des leucémies humaines.

Evi1 (rang 629) est un oncogène dit 'agressif' qui a été trouvé pour la première fois à proximité d'un site d'insertions rétroviraux dans des souris AKXD-23 souffrant de leucémies myéloïdes (Nucifora et al. 2006). Il est au rang 629 parce que l'écart-type entre les échantillons est très grand mais il a un profil parfaitement mégacaryoblastique avec une très forte expression dans Mk1 (tableau 4.8).

Tableau 4.8: Résultat des micropuces pour Evi1

	T1	T2	T3	B1	B2	B3	M	E1	E2	E3	Mk1	Mk2	Mk3
1438325 at	-1,8	-0,9	-1,4	-1,7	-1,7	-1,4	-1,8	-1,1	-1,4	-1,5	7,3	2,4	4,8

Son expression dans les leucémies AML humaines est souvent associée à de faibles chances de survie. Son taux d'expression est globalement bas dans les tissus normaux (Nucifora et al. 2006), ce qui pourrait en faire une bonne cible thérapeutique. De façon très intéressante, certaines études font penser que Evi1 pourrait être associé aux leucémies mégacaryoblastiques et aux mégacaryocytes. Premièrement, la protéine de fusion dans lequel il est impliqué, AML1-Evi1 va préférentiellement provoquer des leucémies AMKL dans des souris transgéniques (Maki et al. 2006). De plus Evi-1, présent dans les mégacaryocytes et plaquettes chez l'humain, aurait un rôle favorable dans l'engagement de la différenciation de la lignée mégacaryocytaire (Shimizu et al. 2002, Kilbey et al. 2005). Donc les résultats des micropuces renforcent l'hypothèse de l'implication de ce gène dans les leucémies AMKL. Une thérapie par du trioxide d'arsenic provoque la dégradation de la protéine de fusion AML1-Evi1 (Shackelford et al. 2006).

Quatre autres gènes parmi les 50 premiers *probesets* pourraient être potentiellement impliqués directement dans la pathologie. Il s'agit de Cbfa2t3h, Vezf, Ifitm1 et Srgap3.

Cbfa2t3h (rang 43) est plus connu sous le nom d'Eto-2. La protéine fait partie d'une famille de co-répresseurs. Comme ETO-1, ETO-2 a été originellement identifiée dans des translocations chromosomiques de leucémies aigues myéloïdes. Chez l'humain, ETO-2 est connue pour être exprimée fortement dans les cellules CD34⁺ et est réprimée lors de la différenciation érythroïde et granulocytaire. Assez récemment, il a été trouvé que Eto-2 est un partenaire d'interaction du complexe Scl dans les cellules érythroïdes et mégacaryocytaires mais le complexe multi-protéique serait légèrement différent dans les 2 types de cellules (Schuh et al. 2005). D'après le résultat des micropuces, Scl, et un de ses partenaires d'interaction Lmo2 sont fortement exprimés dans les leucémies E et Mk, alors que Eto2 est spécifique aux leucémies mégacaryoblastiques seulement. Eto-2 pourrait être un candidat intéressant impliqué dans la déviation soit vers la lignée érythroïde ou mégacaryocytaire à partir du progéniteur bipotent. Il serait intéressant de savoir si l'expression soutenue de Eto-2 a eu un rôle dans l'apparition ou la progression des leucémies Mks induites par le rétrovirus Graffi. Il serait aussi intéressant de connaître les cibles d'Eto-2 à la fois dans les cellules érythroïdes et mégacaryocytaires.

Vezf1 (rang 20) est un facteur de transcription avec des domaines en doigt de zinc qui est impliqué dans l'angiogenèse (Miyashita et al. 2004). Ifitm1 (rang 24) est un récepteur de surface qui activerait les propriétés anti-prolifératives des interférons mais il aurait également un rôle dans l'invasion des cellules tumorales et dans la progression du cancer par suppression des cellules NK (Akyerli et al. 2005). Srgap3 (rang 33) serait impliqué dans la migration cellulaire (réorganisation du cytosquelette) (Yang et al. 2006).

Deux autres gènes, Ffar2 et C3ar1 ont un rôle connu dans l'hématopoïèse mais ne sont pas encore reliés à la lignée mégacaryocytaire. Ffar2 (rang 18) code pour un récepteur couplé aux protéines G. Il est connu pour être exprimé fortement par les neutrophiles et monocytes (confirmé par une expression dans la leucémie myéloïde) et la protéine serait activée par des chaînes courtes d'acides gras (Nilsson et al. 2003). C3ar1 (rang 19) est connu pour être exprimé par plusieurs cellules présentatrices d'antigènes (cellules B, cellules dendritiques, monocytes/macrophages) et l'activation de ce récepteur par l'anaphylatoxine C3a provoque la migration des cellules vers les sites d'infection et agit sur la modulation de la réponse immunitaire (Gutzmer et al. 2006).

D'autres gènes qui ne sont pas encore reliés ni à la lignée mégacaryocytaire ni à la leucémie mégacaryoblastique méritent d'être étudiés Il s'agit par exemple de Fkbp9, Gulp1, Gucyl1a3, St8sia6 (transfert d'acide sialique), Ica1 (sécrétion des granules),

Cuedc1 (domaine d'ubiquitination). Fkbp9, Gulp1, Gucyl1a3, et l'oncogène Kit ont été sélectionnés pour être validés par RT-PCR et leurs caractéristiques vont être décrites dans la section 4.5.2.4.4.

4.5.2.4.2. Gènes surexprimés dans MK1 MK2 MK3 et E1

A l'issue de l'analyse SAM, 892 *probesets* ont passé les critères d'une valeur p inférieure à 0,01 et un FDR inférieur à 0,20 pour les gènes spécifiques de Mk1, Mk2, Mk3 et E1. Le tableau correspondant (appendice C) montre les résultats des 50 premiers *probesets* et est organisé de façon semblable à l'appendice B. Une colonne supplémentaire (E1-E) indique les ratios d'expression des *probesets* entre la leucémie E1 et le contrôle érythroïde. De façon logique, les gènes de ce tableau qui traduisent les caractères mégacaryoblastiques de E1 sont sous-exprimés dans le contrôle érythroïde par rapport à E1.

Les 50 premiers *probesets* correspondent à 44 gènes. Les ratios d'expression varient très significativement, ce qui montre qu'il était très utile de faire ce tableau.

Parmi les 44 gènes, 7 gènes sont très spécifiques de la lignée mégacaryocytaire. Il s'agit de Gp9, Esam1, Itga2, Angpt1, Cxcl4, Cd151, Rab27. Les rôles de Itga2, CD151 et Cxcl4 (Pff4) ont été discutés ci-dessus (sections 4.5.2.3.1. et 4.5.2.3.2). Gp9 code pour une sous-unité du récepteur de Willebrand. Un défaut dans ce gène conduit au syndrome hémorragique de Bernard-Soulier (Lanza et al. 2006). Esam1 est une molécule d'adhésion plaquettaire (Nasdala et al. 2002). Angpt1 qui régule la formation des vaisseaux sanguins est connu pour être exprimé par les mégakaryocytes (Li , Huang et al, 2001). Rab27 appartenant à la famille de l'oncogène Ras, joue un rôle prédominant dans la sécrétion des granules (Tolmachova et al. 2007).

Les autres gènes ne sont pas connus en relation avec la lignée mégacaryocytaire ou les leucémies mégacaryoblastiques et méritent d'être étudiés dans ce sens. Parmi eux, on peut citer le gène Riken, 1700012H17Rik, qui est au premier rang du tableau et dont la fonction est totalement inconnue ainsi que 1110019K23Rik (rang 13), le *probeset* 15 ou 2610204M08Rik (rang 28). Ceci témoigne du fait que la lignée mégacaryocytaire ainsi que les leucémies mégacaryoblastiques sont peu connues.

Parmi les gènes non connus en relation avec la lignée mégacaryocytaire, 3 gènes semblent avoir des propriétés oncogéniques. Il s'agit de *Pard3b* (rang 3), *Dusp3* (rang 19) et *Myct1* (rang 35). *Pard3b* comporte un domaine PDZ et pourrait être impliqué notamment dans le réarrangement des chromosomes lors de la mitose. *Dusp3* est quant à lui un inhibiteur de MAPKinases. La description de sa fonction gravite autour des voies de signalisation impliquées dans la régulation du cycle cellulaire et de la division (Cerignoli et al. 2006). *Myctn1* (*Myc target gene 1*) est comme son nom l'indique, une cible positive de l'oncogène *Myc*. *Myc* a de nombreux rôles et ses actions passent par de nombreux effecteurs. *Myctn1*, en plus d'être régulé par *Myc*, a aussi les effets transformants et tumorigéniques de *Myc* et a donc des propriétés oncogéniques (Rogulski et al. 2005).

Quatre autres gènes semblent avoir un rôle lié à une fonction non-leucémique (fonction hématopoïétique). Il s'agit de *Stxbp5* (rang 8), *Pdzk1ip1* (rangs 13 et 16), *Tbxas1* (rang 17), *Tek* (rang 21).

Stxbp5 n'est pas encore connu pour être exprimé par les mégacaryocytes mais sa fonction est sûrement reliée à la sécrétion des granules plaquettaires car il est impliqué dans la régulation du transport des vésicules et de l'exocytose (Yizhar et al. 2006).

Pdzk1ip1 interagit avec la protéine *Pdzk1*. Une étude montre de façon intéressante que *Pdzk1ip1* se trouve sur le même locus que *Scl* (*Tal1*) et que leurs expressions sont corrélées dans de nombreuses lignées cellulaires testées et que ces 2 gènes partageraient les mêmes séquences régulatrices (Delabesse et al. 2005).

Tbxas1 est responsable de la synthèse de la thromboxane A2 qui induit l'agrégation plaquettaire en se fixant sur le récepteur de la thromboxane A2 (*Tbxa2R*). *Tbxa2R* est effectivement spécifiquement surexprimé dans *Mk1*, *Mk2* et *Mk3* (non-montré). Cependant aucune information dans la littérature n'indique que *Tbxas1* soit exprimé dans la lignée mégacaryocytaire. Le même phénomène est observé avec *Tek* (*Tie-2*) qui est le récepteur à l'angiopoïétine 1 (codée par *Angp1*, rang 22). *Tek* est connu pour être régulé par *Fli-1* qui est surexprimé effectivement dans *E1*, *Mk1*, *Mk2* et *Mk3* (Hart et al. 2000, tableau 4.5). Le récepteur *Tie1* est connu pour être exprimé par les plaquettes mais pas *Tie2* (Tsiamis et al 2000). *Tie 1* est effectivement fortement surexprimé dans *Mk1*, *Mk2* et *Mk3* (non montré). Il serait intéressant de vérifier ces expressions pour voir s'il s'agit

d'une conséquence de l'état tumoral. Ces expressions concomitantes de récepteur et du ligand provoquent des boucles d'autostimulation.

4.5.2.4.3. Gènes surexprimés dans MK1 MK2 MK3, M et/ou B

A l'issue de l'analyse SAM, 1041 *probesets* ont passé les critères d'une valeur p inférieure à 0,01 et un FDR inférieur à 0,20 pour les gènes spécifiques de Mk1, Mk2, Mk3 et M, et 1281 *probesets* ont passé les critères pour Mk1, Mk2, Mk3, M, B1, B2, B3. Le tableau correspondant (appendice D) montre les 50 premiers *probesets* spécifiques des échantillons Mk et M et l'appendice E montre les 50 premiers *probesets* spécifiques des échantillons Mk et M et B. Parmi ces gènes, certains étaient déjà connus dans la lignée myéloïde (9 gènes dans chaque tableau), certains étaient connus pour être exprimés dans les lignées myéloïde et B (4 dans l'appendice E), ou dans les lignées myéloïde et mégacaryocytaire (1 dans l'appendice D, 2 dans l'appendice E) et 3 gènes sont connus pour être exprimés dans les 3 lignées, myéloïde, B et mégacaryocytaire (appendice E).

4.5.2.4.3.1. Gènes impliqués dans le cancer

Parmi les gènes spécifiques de ces 2 tableaux, 12 suscitent de l'intérêt de par leur relation très probable avec le processus de cancérisation. Il s'agit de:

- 1) FosL2 (4.11, rang 21): il est un des 4 membres de la famille Ras et interagit avec Jun (4.12, rang 181, non montré) pour former le facteur de transcription AP-1.
- 2) Rbpsuh (Rbpj, 4.11 rang 48): ce gène est important pour la mise en place de l'hématopoïèse embryonnaire et se situe directement en aval de Notch (Robert-Morena et al 2005).
- 3) Ctbp2 (4.11 rangs 3 et 15): est une molécule co-répresseur qui interagit notamment avec Evil et Fog.
- 4) Kit (4.11, rangs 18 et 23): la description de cet oncogène se retrouve dans la section 4.5.2.4.5 (gènes sélectionnées pour validation RT-PCR)

- 5) Itga5 (4.11 rangs 31, 3 *probesets*): il est le récepteur de la fibronectine et joue également un rôle dans la migration des cellules tumorales et son interaction avec Tat de VIH serait impliquée dans le syndrome de Kaposi.
- 6) Pld1 (4.11 rangs 34, 3 *probesets*): la phospholipase D est impliquée dans la signalisation grâce à la production de seconds messagers. Elle est connue en relation avec les activités oncogéniques de Ras.
- 7) Ptpre (4.11 rangs 38, 3 *probesets*): est également une molécule de signalisation impliquée dans la signalisation de Ras, la mitose, la différenciation.
- 8) S110a4 (4.11 rang 44): est impliquée dans les métastases (Garett et al. 2006)
- 9) Sfp1 (4.12 rang 16): ce gène n'est autre que Spi1 (Pu.1), facteur de transcription impliqué dans la différenciation de la lignée B et des cellules myéloïdes (voir chapitre I).
- 10) Fes (4.12 rangs 12 et 18): cet oncogène est une tyrosine kinase. Il a été retrouvé dans des translocations chromosomiques de patients souffrant de APL (acute promyelocytic leukemia). Il est connu pour être présent dans la lignée myéloïde (Yates and Gasson, 1996), la lignée B (DeCino et al. 1988) et joue également un rôle dans l'agrégation plaquettaire (Senis et al. 2003).
- 11) Tpd52 (4.12 rangs 22, 3 *probesets*): cet oncogène est connu pour initier la transformation cellulaire (Lewis et al. 2007).
- 12) Rassf3 (4.12 rangs 22, 3 *probesets*): membre de la famille Ras.

En résumé des tableaux des gènes spécifiques des leucémies mégacaryoblastiques induites par le rétrovirus Graffi, certains oncogènes connus ont été rencontrés et leur l'expression soutenue pourrait être impliquée dans le blocage de la différenciation. Il s'agit de Pu.1, Kit, Gata2 (et 1), Fli-1, Fog-1, Nfe2, AML-1(Runx1), Scl et ses partenaires d'interactions (Lmo2, Eto2).

4.5.2.4.3.2. Gènes probablement impliqués dans la fonction normale des mégacaryocytes/plaquettes

Dans les tableaux 4.11 et 4.12, 2 gènes seulement sont connus pour être associés à la fonction mégacaryocytaire. Il s'agit de Scin (appendice E rang 33) et Ptgs1 (appendice D rangs 29 et 36). Scin (scinderin) est similaire à la gelsoline (appendice E, rangs > à 50, 5 *probesets*, non montré). Ces 2 protéines sont impliquées dans la réorganisation du cytosquelette lors de l'exocytose. PTGS1 (Cox1) est impliquée dans l'agrégation plaquettaire en participant à la formation de la thromboxane A2 et l'aspirine a des propriétés fluidifiantes du sang car elle inhibe Cox1 plaquettaire.

En analysant les résultats et au-delà des 50 premiers *probesets* indiqués sur les 2 tableaux, il ressort de façon évidente que les gènes communs des leucémies M et B et Mk semblent partager une fonction commune. Ces gènes sont fréquemment connus pour être exprimés par les cellules B, les macrophages ou les neutrophiles, et par les cellules dendritiques.

Le partage d'une fonction commune se voit notamment dans la répartition de la distribution cellulaire des gènes, avec une forte proportion des gènes codant pour des récepteurs membranaires (proportion plus élevée que pour les autres tableaux analysés, non montrés). La fonction commune révélée ici est celle reliée à l'immunité innée voire même adaptative, et aux fonctions des cellules présentatrices d'antigène.

Ceci peut paraître surprenant à première vue car les plaquettes, petites cellules énuclées, sont surtout connues pour être impliquées dans la coagulation (sécrétion de facteurs de coagulation et formation du caillot par agrégation). Les plaquettes sécrètent beaucoup plus que quelques facteurs de coagulation: le secrétome des plaquettes est bien caractérisé. Il est composé des granules α , des granules denses et des lysosomes. Lors de l'activation plaquettaire, tout ce qui se trouve dans ces organelles va être sécrété à l'extérieur. Les granules α contiennent notamment les protéines d'adhésion et de réparation (facteurs de Willebrand, fibrinogène, thrombospondine,...). Les granules denses contiennent des petites molécules (sérotonine –vasoconstricteur-, calcium, ADP) et les lysosomes contiennent des enzymes permettant de défaire les agrégats plaquettaires. En relation avec cela, les analyses des micropuces mettent fréquemment en évidence des gènes impliqués dans le trafic vésiculaire et le réarrangement du cytosquelette.

La fonction commune aux leucémies B, M et Mk est celle d'une implication plus complexe dans l'immunité innée et adaptative. Cette implication plaquettaire et mégacaryocytaire, restée longtemps ignorée, tend à être de plus en plus reconnue (Weyrich et Zimmerman, 2004, Elzey et al. 2005). Le tableau correspondant (appendice F) explique ces résultats en montrant la répartition des gènes communs aux leucémies Mk, M et/ou B selon leurs fonctions. Les gènes inclus dans cette analyse ne sont pas tous compris dans les tableaux 4.11 et 4.12 car ils peuvent se situer au-delà des 50 premiers *probesets* présentés. Seuls les gènes avec une fonction liée à la réponse immunitaire innée ou adaptative ont été sélectionnés pour composer l'appendice F.

Le tableau de l'appendice F se divise en 2 grandes parties, la première correspondant aux protéines sécrétées et l'autre aux protéines membranaires et cytoplasmiques. La fonction des gènes est résumée dans la 4^{ème} colonne.

Parmi les facteurs sécrétés (ou extracellulaires) se trouvent des granules antimicrobiens et des cytokines. Les granules antimicrobiens peuvent tuer les pathogènes et les cytokines (interleukines et chimiokines) et peuvent stimuler la réponse immunitaire. Les cathepsines sont des protéases activées par une baisse du pH des lysosomes. Elles peuvent notamment être impliquées dans l'apprêt des antigènes ingérés. La cathepsine G est connue pour provoquer la formation de prothrombinase et de fibrine.

Parmi les gènes codant pour des protéines membranaires ou cytoplasmiques, plusieurs sections sont présentées: 1) Inflammation, LPS, microbe; 2) Récepteurs Toll-like; 3) Récepteurs Fc; 4) Récepteurs modulateurs de la fonction immunitaire; 5) la présentation des antigènes.

1) Inflammation, LPS, microbe:

Les cellules de l'immunité innée déclenchent le processus de l'inflammation en réponse aux stimuli des pathogènes. Le LPS situé au niveau des parois bactériennes est un puissant stimulus auquel réagissent par exemple les gènes *Lst1* et *Litaf*.

2) Récepteurs Toll-like et protéines associées:

Les récepteurs Toll-like font partie des récepteurs de la réponse immunitaire innée: ils reçoivent les stimuli microbiens (bactéries, virus, champignons), déclenchent l'inflammation et font le relais entre l'immunité innée et adaptative (activation lymphocytaire). Des gènes codant pour les récepteurs Toll-like ou codant pour des

molécules de signalisation associées sont surexprimées dans les leucémies mégacaryoblastiques induites par le rétrovirus Graffi.

Récemment l'expression du Toll-like récepteur 4 (Tlr4) a été décrite à la surface des plaquettes même si son expression est assez faible d'après les résultats des micropuces. Afin d'activer le TLR4, le LPS doit se combiner à une molécule de LBP (LPS binding protein) et au CD14 (appendice F). L'activation du TLR4 envoie des signaux intracellulaires en passant notamment par la protéine Myd88 et NfκB pour stimuler les gènes impliqués dans la réponse immunitaire. Il a été montré que l'activation du TLR4 par les plaquettes permettrait le recrutement et l'activation des neutrophiles autour des plaquettes (Clark et al. 2007).

D'après les résultats des micropuces, Tlr2 est le plus fortement exprimé par les échantillons leucémiques mégacaryoblastiques. Il a aussi été détecté sur les plaquettes humaines bien que moins étudié que Tlr4 (Cognasse et al. 2005) mais Tlr2 et Tlr4 sont connus pour agir en synergie contre les bactéries.

SYK est une protéine importante dans la signalisation des Toll-like récepteurs. De plus elle est aussi activée par DECTIN-1 (Clec7a, Récepteurs modulateurs de la fonction immunitaire). SYK intervient aussi dans la signalisation de l'intégrine $\alpha_{IIb}\beta_3$ (section 4.5.2.3.2). Ainsi cette protéine joue un rôle clé dans les plaquettes: activation, réponse immunitaire innée et adaptative.

L'expression de Ebi-3 est augmentée lors de l'activation des récepteurs Toll-like et Irf5 est en aval de la voie de signalisation du Tlr7.

2) Récepteurs Fc:

Les récepteurs Fc et les récepteurs *scavenger* ont comme rôle de permettre l'internalisation des pathogènes dans les cellules phagocytaires (figure 4.9). Les Fc récepteurs reconnaissent des molécules d'anticorps qui sont accrochées aux pathogènes. L'opsonisation des pathogènes par les molécules du complément ainsi que la reconnaissance par les récepteurs du complément facilitent l'internalisation (figure 4.9). L'internalisation du pathogène conduit à la formation du phagosome avec activation de la NADPH oxydase. Les pathogènes sont digérés et les peptides pathogéniques sont apprêtés et présentés aux autres cellules comme antigènes. La protéine SYK est impliquée dans le phénomène de phagocytose des macrophages.

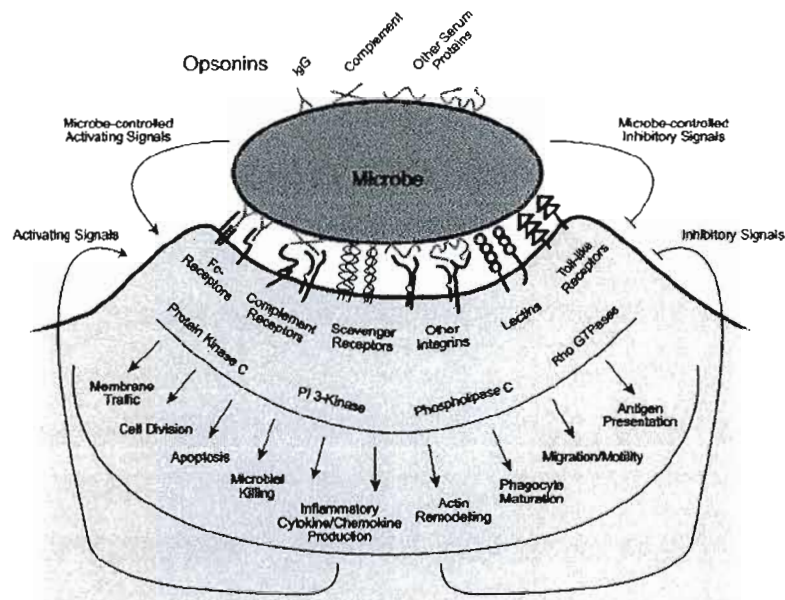


Figure 4.9: Récepteurs et signalisation impliqués dans la phagocytose d'un microbe. Figure tirée de Underhill et Ozinsky, 2002. Les différentes classes de récepteurs présents à la surface d'une cellule phagocytaire sont illustrés.

2) Récepteurs modulateurs de la réponse immunitaire.

Dans cette section, se trouvent les récepteurs lectine de type C qui représentent la 3ème catégorie de récepteurs de la réponse innée avec les récepteurs Toll-like et les récepteurs *scavenger*. Ils sont impliqués dans l'internalisation et la médiation de l'inflammation (figure 4.9). Clec7a (Dectin-1) est spécialisée dans la détection des levures et l'internalisation. Dectin-1 agirait en synergie avec TLR2 pour augmenter la réponse pro-inflammatoire (cytokines) (Gantner et al. 2003).

D'autres récepteurs sont inclus dans cette section. Ils modulent l'activité de la cellule cible comme par exemple les récepteur des interleukines et du TNF.

2) Présentation d'antigène.

Ces gènes ne sont pas connus pour être exprimés par les plaquettes. Les molécules du CMH II sont présentes à la surface des cellules présentatrices d'antigènes, comprenant les macrophages, les cellules dendritiques et les cellules B. L'expression de ces gènes est étonnante car ce n'est pas connu pour la lignée mégacaryocytaire et pourtant cela est en

corrélation avec la description des fonctions des gènes ci-dessus (réponse immunitaire innée). La présentation des antigènes exogènes se fait après la phagocytose des pathogènes. Les récepteurs et autres protéines nécessaires à la phagocytose sont largement représentés dans l'appendice F. De plus, des phénomènes de phagocytose sont connus pour les plaquettes et les mégacaryocytes. Pour les mégacaryocytes, cellules géantes, ce phénomène se nomme emperipolèse (*emperipolysis*) (figure 4.10 A): le mégacaryocyte avale des cellules entières qui peuvent ressortir intactes de l'autre côté de la cellule. Le mégacaryocyte est une cellule rare et peu étudiée et le phénomène d'emperipolèse a été observé souvent lors de cas pathologiques. De façon intéressante, des observations ont montré que ce phénomène s'amplifie lors de l'activation des mégacaryocytes par le LPS et qu'il y avait augmentation de la molécule ICAM-1 (appendice F) aux bordures des membranes phagocytantes (Tanaka et al., 1997). La phagocytose des plaquettes concernent de petites bactéries et virus et se fait grâce au récepteur DC-SIGN, récepteur de type lectine C très similaire à DECTIN-1 (figure 4.10 B, Youssefian et al., 2002).

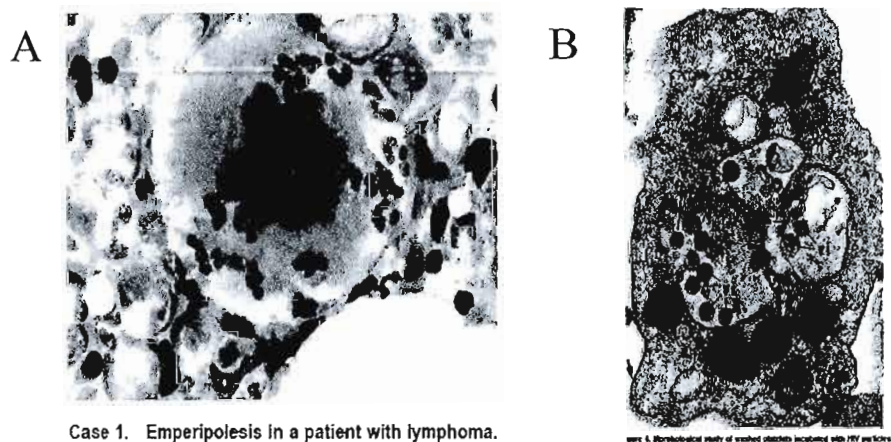


Figure 4.10: Phagocytose par les mégacaryocytes et les plaquettes. (A) Empérilopèse de neutrophiles anormaux par les plaquettes, Samii et Pasteur, 1998 (B) Phagocytose de particules virales par une plaquette, Youssefian et al. 2002.

En conclusion, ces éléments ne concernent pas la leucémogénèse. De plus en plus d'études vont dans le sens d'un rôle beaucoup plus important des plaquettes dans la réponse immunitaire influençant la réponse innée et adaptative. D'ailleurs, certains chercheurs réclament que ces cellules soient classifiées comme cellules effectrices du

système immunitaire (Weyrich et Zimmerman, 2004, Elzey et al. 2005, Sprague et al. 2005, Cognasse et al. 2005).

Les résultats proviennent d'échantillons leucémiques et donc les résultats doivent être validés sur des échantillons non-leucémiques avant que toute hypothèse soit émise. Ces gènes sont ressortis de façon évidente de l'analyse. Ils étaient visibles dès l'étude du profil général des résultats des micropuces et cette analyse ne pouvait pas être écartée. Il serait très stimulant de vérifier l'expression de ces gènes dans des mégacaryocytes et plaquettes humaines et murines. Le mégacaryocyte est une cellule géante et mystérieuse. On pourrait supposer qu'il a des fonctions additionnelles par rapport aux plaquettes énuclées.

La figure 4.11 résume les gènes trouvés exprimés par les leucémies mégacaryoblastiques induites par le rétrovirus Graffi et impliqués dans la réponse immunitaire.

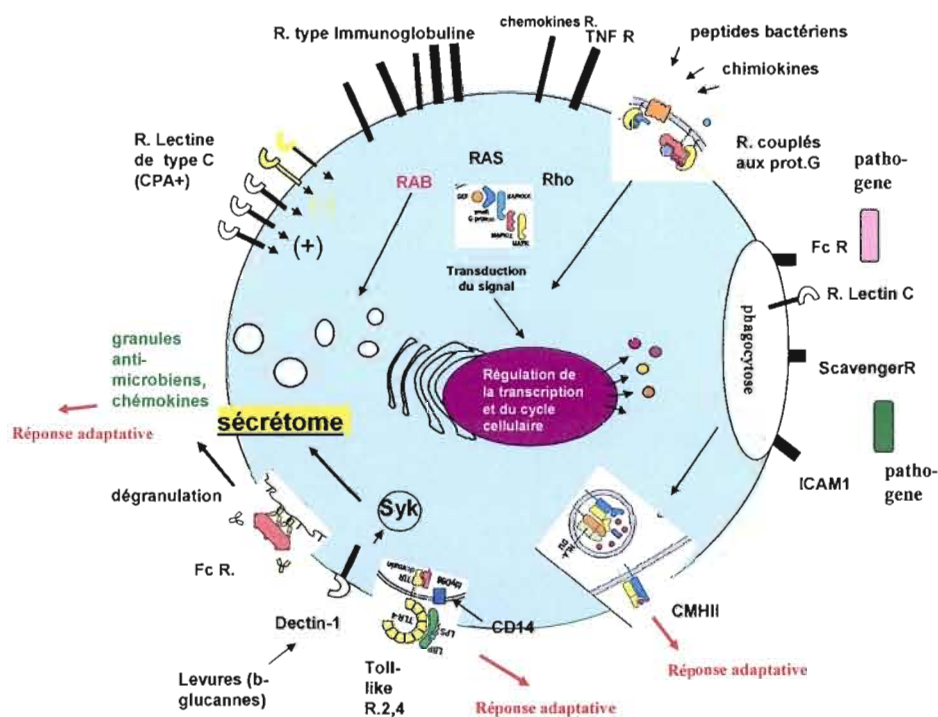


Figure 4.11: Schéma des gènes impliqués dans la réponse immunitaire exprimés par les cellules mégacaryoblastiques induites par le rétrovirus Graffi. Le schéma représente une cellule leucémique mégacaryoblastique induite par le rétrovirus Graffi. Les protéines correspondant à des gènes spécifiquement surexprimés dans ces leucémies et impliqués dans la réponse immunitaire sont illustrées.

4.5.2.4.5. Gènes mégacaryoblastiques sélectionnés pour les tests de RT-PCR

Certains gènes spécifiques des leucémies mégacaryoblastiques ont été sélectionnés pour être validés en RT-PCR avec d'autres échantillons. Les résultats de ces tests sont montrés plus loin dans la section 4.6.3 mais les caractéristiques des gènes sélectionnés sont expliquées ici. Ces gènes sont *Gulp1*, *Fkbp9*, *Gucy1a3*, *Ncf2*, *Gata2* et *Kit*. Le tableau 4.9 présente les résultats des micropuces (valeurs de SAM et valeurs de ratios, *probesets* redondants) pour ces gènes. La première colonne du tableau indique à partir de quel tableau ils ont été sélectionnés. Le tableau indique également les rangs occupés par ces gènes dans ces tableaux (6ème colonne). Le tableau indique également le nom des gènes, le score (SAM d-score), la valeur p, la valeur FDR, le nombre de *probesets* correspondant à chaque gène sélectionné ainsi que les ratios d'expression.

Tableau 4.9: Gènes mégacaryocytaires sélectionnés pour l'analyse par RT-PCR

tableaux	noms	SAM d-score	P-value	FDR	rangs	probesets	T1	T2	T3	B1	B2	B3	M	E1	E2	E3	Mk1	Mk2	Mk3
4.8 (Mk)	<i>Gulp1</i>	14,2	<0,001	0,07	16	1/3	-0,7	-0,7	-0,5	-0,6	-0,9	-0,9	-0,9	-1,0	-0,9	-0,9	3,1	2,6	2,3
4.8 (Mk)	<i>Gulp1</i>	11,2	<0,001	0,07	36	2/3	-0,6	-0,5	-0,8	-0,6	-0,6	-0,8	-0,7	-0,7	-0,5	-0,5	1,7	2,7	2,1
4.8 (Mk)	<i>Gulp1</i>	4,9	0,003	0,13	466	3/3	-0,4	-0,3	-0,5	-0,4	-0,1	-0,3	-0,2	-0,3	-0,2	-0,2	1,4	0,6	0,8
4.8 (Mk)	<i>Fkbp9</i>	23,7	<0,001	0,07	1	1/2	-1,4	-1,3	-1,6	-1,3	-1,3	-1,3	-1,1	-1,0	-1,0	-1,1	4,1	4,6	3,8
4.8(Mk)	<i>Fkbp9</i>	10,1	<0,001	0,07	60	2/2	-1,0	-0,7	-1,0	-0,6	-0,8	-0,6	-0,9	-0,6	-0,4	-0,6	2,3	2,8	2,8
4.8 (Mk)	<i>Gucy1a3</i>	18,2	<0,001	0,07	6	1/3	-1,2	-0,6	-1,0	-1,0	-1,1	-1,2	-1,0	0,2	-1,1	-1,0	3,0	3,2	2,7
4.8 (Mk)	<i>Gucy1a3</i>	10,8	<0,001	0,07	47	2/3	-0,7	-0,8	-0,5	-0,7	-0,6	-0,5	-0,4	-0,1	-0,6	-0,4	2,2	1,5	1,6
4.8 (Mk)	<i>Gucy1a3</i>	7,8	<0,001	0,08	113	3/3	-0,3	-0,4	-0,2	-0,3	-0,3	-0,2	-0,2	0,1	-0,3	-0,1	0,5	0,8	1,1
4.12 (Mk,M,B)	<i>Ncf2</i>	18,5	<0,001	<0,01	1	1/2	-2,6	-2,4	-2,1	2,0	1,6	1,6	3,0	-2,1	-2,2	-2,4	1,6	2,5	1,6
4.11 (Mk, M)	<i>Kit</i>	10,5	<0,001	0,01	18	1/3	-2,2	-2,1	-2,2	-2,2	-2,3	1,2	2,8	-1,7	-1,6	-2,2	3,8	4,6	4,0
4.11 (Mk, M)	<i>Kit</i>	10,0	<0,001	0,01	23	2/3	-1,5	-1,5	-1,3	-1,6	-1,5	0,0	2,1	-1,0	-1,1	-1,4	2,7	3,8	2,3
4.11 (Mk, M)	<i>Kit</i>	5,5	<0,001	0,04	195	3/3	-0,6	-0,4	-0,7	-0,1	-0,5	-0,2	0,3	-0,6	-0,5	-0,3	0,9	1,4	1,4
4.15 (Mk, E)	<i>Gata2</i>	7,4	<0,001	0,01	115	1/2	-2,3	-2,0	-2,3	-2,1	-2,4	-1,9	-1,6	0,2	2,4	1,1	3,6	4,0	3,4
4.15 (Mk, E)	<i>Gata2</i>	5,4	<0,001	0,02	317	2/2	-1,9	-1,6	-1,9	-1,7	-1,8	-1,9	-1,1	-0,6	2,2	0,6	3,4	3,4	2,8

Gulp1, *Fkbp9*, *Gucy1a3* sont spécifiques uniquement des leucémies mégacaryoblastiques, *Ncf2* est commun aux leucémies Mk, B et M. *Kit* est exprimé communément par les leucémies Mk et M, et *Gata2* est exprimé communément par les leucémies E et Mk avec une expression plus forte dans les leucémies Mk.

4.5.2.4.4.1. Gulp1 (engulfment adaptor PTB domain containing 1)

Gulp1 est situé sur le chromosome 1 chez la souris et les protéines murine et humaine partagent 92% d'homologie.

Gulp1 (Ced-6) a premièrement été découvert chez *C. elegans* appartenant à une famille de gènes (Ced-1 à Ced-7) impliqués dans la phagocytose des cellules apoptotiques. Ced-1, connu également sous le nom de Lrp1 (low density lipoprotein-receptor related protein 1) code pour un récepteur à plusieurs ligands qui joue un rôle dans la signalisation et la prise en charge de lipoprotéines. La protéine GULP1 est une molécule adaptatrice de LRP1, qui permettrait de réguler le trafic intracellulaire de LRP1 (Kiss et al. 2006). Ainsi Gulp1 semble impliqué dans l'endocytose des récepteurs, le transport du cholestérol et des sphingolipides à travers les endosomes tardifs. Plus précisément, Gulp1 semble impliqué dans le transport de la prosaponine (Kiss et al. 2006). Les saponines sont importantes pour la dégradation lysosomiale et également pour la présentation des antigènes lipidiques aux cellules T (Winau et al. 2004). Ainsi, on peut penser que Gulp1 a un rôle très important dans les mégacaryoblastes ou mégacaryocytes. On sait que les mégacaryocytes sont composés de structures lipidiques et d'un système de sécrétion (le sécrétome) complexe. Les résultats des micropuces montrent de fortes surexpressions de Lrp1 dans 2 leucémies Mk (MK2 et MK3, tableau 4.10), ceci étant en accord avec le rôle de molécule adaptatrice de Gulp1 pour Lrp1.

Tableau 4.10: expression de Lrp1

	T1	T2	T3	B1	B2	B3	M	E1	E2	E3	Mk1	Mk2	Mk3
Lrp1	-0,04	-0,57	-1,01	-0,86	-1,17	-0,64	2,30	-0,02	-0,19	-0,49	-0,45	0,91	2,25

On connaît peu de chose à propos de Gulp1 et sa présence n'a pas encore été détectée dans les mégacaryoblastes et mégacaryocytes. L'étude des ligands associés à Lrp1 et Gulp1 dans la lignée mégacaryoblastique est donc très intéressante.

4.5.2.4.4.2. Fkbp9 (FK506 binding protein 9)

Le gène est localisé sur le chromosome 6 et les protéines murine et humaine ont 53% d'homologie.

Fkbp9 a été isolé à partir du thymus d'une souris SCID (Jo et al. 2001). Il appartient à la famille des protéines liant FK506 (FKBP, FK506 binding proteins, immunophilines). La molécule FK506 est un immunosuppresseur et la rapamycine est structuellement proche de FK506. La rapamycine est aussi un immunosuppresseur capable de se lier aux protéines de la famille FKBP même si elle est connue pour agir sur la voie de signalisation Akt passant par le récepteur mTor. Il a été montré que la rapamycine a également un effet pro-apoptotique sur des cellules leucémiques aiguës lymphoblastiques et que cet effet est exercé de façon indépendante de mTor (Avellino et al. 2005). Cette étude suggère que la rapamycine pourrait être efficace pour traiter les leucémies lymphoblastiques aiguës de l'enfant (Avellino et al. 2005).

Comme Gulp1, très peu de choses sont connues sur Fkbp9 et donc il est très intéressant de poursuivre son étude. Il serait intéressant de connaître les ligands de Fkbp9, en testant par exemple l'effet de la rapamycine sur la survie des cellules leucémiques mégacaryoblastiques.

4.5.2.4.4.3. Gucy1a3 (Guanylate cyclase 1, soluble, alpha 3)

Le gène Gucy1a3 est localisé sur le chromosome 3 chez la souris et les protéines murine et humaine partagent 82% d'homologie.

Gucy1a3 fait partie de la famille des guanylates cycliques solubles. Ce sont des molécules qui contiennent un hème et qui répondent au signal d'activation fourni par l'oxyde nitrique (NO) (Poulos T, 2006). Elles sont généralement cytosoliques mais peuvent s'ancrer à la membrane lorsqu'il y a un signal d'activation fourni par une élévation de calcium intracellulaire. Quand les guanylates cyclases sont activées, leurs activités GTP cyclases augmentent et des molécules de GMPc sont alors libérées dans le cytosol et vont agir sur plusieurs voies de régulation (Poulos T, 2006). Elles fonctionnent en hétérodimère avec une molécule α et une molécule β (Poulos T, 2006). Ainsi, d'après

le résultat des micropuces, Gucy1b3 (tableau 4.11) est bien également spécifique des leucémies mégacaryoblastiques.

Tableau 4.11: Expression de Gucy1b3

	T1	T2	T3	B1	B2	B3	M	E1	E2	E3	Mk1	Mk2	Mk3
Gucy1b3	-0,50	-0,47	-0,20	-0,69	-0,37	-0,49	-0,57	-0,08	-0,55	-0,34	1,90	0,51	1,84
Gucy1b3	-0,58	-0,73	-0,87	-0,41	-0,65	-0,64	-0,73	-0,45	-0,50	-0,48	2,54	0,77	2,73

On sait aussi que l'expression génique des guanylates cyclases est modifiée dans des cas d'hypoxie, d'hypertension, de réponse inflammatoire et lors du vieillissement. Cependant, le mécanisme de la régulation reste inconnu (Poulos T, 2006). Dans une étude, Gucy1a3 est surexprimé dans des tissus de cancer de la prostate par rapport à des tissus non cancéreux (Dong et al. 2005). Une autre étude suggère que le NO et Gucy1a3 et b3 agissent sur l'expression de Vegf, sur-exprimé dans des gliomes malins, favorisant la vascularisation des tumeurs (Saino et al. 2004).

Est-ce que le NO, qui active les guanylates cyclases, joue un rôle connu dans la signalisation mégacaryocytaire ou plaquettaire? Des études ont montré que NO semble effectivement avoir un effet sur l'activation des plaquettes. En fait son rôle semble biphasique; de faibles concentrations de NO activent les plaquettes et de fortes concentrations inhibent l'activation (Marjavonic et al. 2005). Le NO semble également jouer un rôle dans les mégacaryocytes. Une étude récente montre que le NO favorise la formation des plaquettes par les mégacaryocytes pro-plaquettaires et entraîne l'apoptose de ces mêmes mégacaryocytes (Battinelli et al. 2001). Ainsi, d'après les résultats des micropuces, on peut émettre l'hypothèse que les effets du NO passent spécifiquement par l'hétérodimère Gucy1a3/b3 dans les cellules de la lignée mégacaryocytaire. Il serait très intéressant d'en poursuivre l'étude fonctionnelle.

4.5.2.4.4.3. Ncf2 neutrophil cytosolic factor 2

Le gène Ncf2 est situé sur le chromosome 1 chez la souris. Les protéines murine et humaine partagent 84% d'homologie.

Le gène Ncf2 (p67phox) est une sous-unité de la NADPH oxydase (Sheppard et al. 2005). La NADPH oxydase fait partie de l'arsenal anti-microbien utilisé par les neutrophiles. Elle produit des molécules d'anion super-oxyde O_2^- . Les anions super-oxydes

sont ensuite transformés en radicaux libres. Une grande partie va être transformée en H_2O_2 par la myéloperoxydase. Plusieurs sous-unités composent la NADPH oxydase. Il s'agit de: Ncf2 (p67phox), Ncf1 (p47phox), Ncf4 (p40phox), Cybb (gp91phox) et Cyba (p22phox) (Sheppard et al. 2005). Il est donc intéressant de voir si les leucémies mégacaryocytaires expriment toutes les sous-unités de la NADPH oxydase (tableau 4.12).

Tableau 4.12: la NADPH oxydase

	T1	T2	T3	B1	B2	B3	M	E1	E2	E3	Mk1	Mk2	Mk3
Ncf2	-2,60	-2,37	-2,09	1,99	1,61	1,61	2,99	-2,14	-2,17	-2,43	1,56	2,46	1,58
Ncf1	-1,80	-2,16	-2,13	-0,58	-0,89	-1,73	3,58	-0,21	-1,98	-2,51	2,73	5,00	2,69
Ncf1	-1,75	-1,80	-1,51	-0,51	-0,77	-1,64	3,50	-0,99	-1,85	-2,14	2,52	4,99	1,95
Ncf4	-2,01	-3,63	1,53	2,01	1,90	1,47	3,26	-2,56	-2,71	-3,60	1,31	1,71	1,33
Cybb	-2,48	-2,81	-2,37	0,36	0,41	2,19	4,02	-0,35	-2,01	-2,86	-0,52	3,44	2,97
Cybb	-2,05	-1,57	-1,93	0,22	0,05	1,48	3,79	-0,63	-1,56	-1,66	-1,76	3,06	2,58
Cyba	-0,38	-0,80	-0,85	-0,68	-0,36	-0,94	1,10	1,82	0,50	0,92	-0,58	-0,54	0,79

Le tableau 4.12 montre qu'effectivement, Ncf1, Ncf4, Cybb et Cyba sont également fortement exprimés dans les leucémies mégacaryoblastiques et dans la leucémie myéloïde induites par le rétrovirus Graffi. Seuls Ncf2, Ncf4 et à moindre degré Cybb sont exprimés par les leucémies de type B. Il est à noter aussi que les leucémies mégacaryoblastiques induites par le rétrovirus Graffi montrent une très forte expression de la myéloperoxydase bien qu'à un degré inférieur à celui de la leucémie myéloïde (non montré). Cette expression est en accord avec la présence de la NADPH oxydase dans les cellules leucémiques mégacaryoblastiques.

La présence de la NADPH oxydase dans les plaquettes a été montrée (Begonja et al.2006). Selon le concept actuel, les molécules d'oxygène réactif joueraient le rôle de seconds messagers et seraient impliquées dans plusieurs fonctions et voies de signalisation dont l'activation plaquettaire. Un modèle très intéressant montre que les molécules d'oxygène réactif seraient transformées en NO par une NO synthase qui activerait à son tour les guanylates solubles (Begonja et al.2006). Ce modèle est tout à fait cohérent avec nos observations d'une forte expression de Gucyl1a3/b3 (voir précédemment). Les molécules de GMPc alors produites agiraient sur d'autres effecteurs pour notamment induire l'activation plaquettaire (Begonja et al.2006).

Beaucoup de chaînons sont encore manquants dans les voies de signalisation impliquant la NADPH oxydase et les guanylates cyclases dans les plaquettes. On voit que

le modèle Graffi avec une analyse détaillée des micropuces pourrait permettre d'élucider plus clairement ces mécanismes.

4.5.2.4.4.4 Gata2, Gata1, Kit

Les fonctions de Gata1, Kit et Gata2 ont été partiellement décrites dans le chapitre I. Gata1 et Gata2 sont 2 facteurs de transcription qui se ressemblent beaucoup. Certaines études montrent un rôle chevauchant et interchangeable pour ces 2 facteurs notamment dans les étapes précoces du développement mégacaryocytaire (Chang et al. 2002).

Dans des souris knock-out pour Gata1, les progéniteurs érythroïdes montrent une expression plus élevée de Gata2, suggérant que Gata2 peut se substituer à Gata1. Cependant, la substitution est partielle puisque la différenciation de ces cellules de souris knock-out est tout de même incomplète (Pevny et al. 1991).

Gata1 est exprimé le plus fortement dans les cellules érythroïdes, mégacaryocytaires et mastocytaires. Gata2 est exprimé dans les érythroblastes, les mégacaryocytes et les mastocytes. Il est également exprimé dans les étapes très précoces de l'hématopoïèse dans lesquelles Gata1 est absent.

Une étude montre que l'expression de Gata2 diminue et doit diminuer pour que la différenciation des cellules érythroïdes ait lieu (Ikonomi et al. 2000). De plus, cette même étude a montré que l'induction d'une expression forte et stable de Gata2 dans les cellules K562 (phénotype érythro-mégacaryocytaire), déséquilibrant le ratio Gata1/Gata2 existant, induit un phénotype beaucoup plus mégacaryocytaire (Ikonomi et al. 2000). Donc si une expression élevée de Gata2 est nécessaire pour la mégacaryopoïèse, cette même expression élevée peut bloquer l'érythropoïèse. Par contre, Gata1 est exprimé dans les mégacaryocytes (Martin et al. 1990). Dans des cellules myéloïdes 416B, qui peuvent se différencier *in vivo* en mégacaryocytes et granulocytes, une surexpression de Gata2 entraîne une différenciation mégacaryocytaire corrélée avec une augmentation de l'expression de Gata1 (Visvader et al. 1993).

En conclusion, le taux relatif d'expression à la fois de Gata1 et de Gata2 à différentes étapes de l'hématopoïèse est critique pour la régulation précise de la différenciation en différentes lignées (Ikonomi et al. 2000).

Il est donc important de bien valider le taux d'expression de Gata1 et Gata2 dans les leucémies induites par le rétrovirus Graffi. Le niveau d'expression de Gata2 dans les leucémies érythroïdes atteste du blocage de ces cellules à un stade encore précoce.

Le proto-oncogène Kit est un récepteur tyrosine kinase. Kit a été découvert comme l'oncogène responsable des activités transformantes du virus du sarcome félin Hardy-Zuckerman IV (Ashman et al. 1999).

L'activation de Kit par son ligand (*stem cell factor*) augmente la survie des progéniteurs mégacaryoblastiques et joue un rôle dans la mégacaryopoïèse (Antonchuk et al. 2004, Zeuner et al. 2007). Il semble que l'activation de Kit soit impliquée dans les stades précoces de la mégacaryopoïèse et peu ou pas impliquée dans les stades tardifs (pas d'effet sur la polyploïdisation ni sur la thrombopoïèse), mais Kit aurait également un effet sur l'agrégation plaquettaire (Avraham et al. 1999). Les actions de Kit sur les progéniteurs mégacaryocytaires se feraient par l'activation du ligand de Kit présent sous forme liée à la surface des cellules stromales de la moelle osseuse et l'effet sur l'agrégation plaquettaire se ferait par les formes solubles du ligand de Kit (Avraham et al. 1999). Beaucoup de signaux vont être envoyés dans la cellule par l'activation du récepteur Kit. Certaines molécules sont connues pour interagir avec les récepteurs tyrosine kinase. Il peut s'agir des PI3 kinases, des phospholipases C-gamma, des kinases de la familles Src ou des protéines activant les GTPases (Avraham et al. 1999). Kit est également impliqué dans de nombreux autres cancers et des inhibiteurs de c-Kit ont été développés mais ces thérapies entraînent comme effets secondaires une thrombopénie sévère (Lengyel et al. 2007, Zeuner et al. 2007). Il serait très intéressant d'élucider précisément les voies de signalisation passant par Kit dans les leucémies mégacaryoblastiques induites par le rétrovirus Graffi en s'aidant premièrement des résultats des micropuces. Une comparaison avec les voies de signalisation impliquées dans les autres cancers pourraient permettre d'établir des stratégies différentes 1) pour traiter les leucémies mégacaryoblastiques 2) traiter les autres cancers Kit⁺ sans effets secondaires sur les plaquettes).

4.5.2.5. Gènes spécifiques des leucémies érythroïdes

De la même façon que pour les leucémies mégacaryoblastiques, une analyse SAM a été réalisée pour identifier les gènes spécifiques des leucémies érythroïdes (section 4.5.2.5.1). Une seconde analyse a été également réalisée pour trouver les gènes spécifiques des leucémies E1, E2 et E3 et également Mk3 (section 4.5.2.5.2), car certaines caractéristiques érythroïdes de Mk3 ont été identifiées par l'établissement du profil général des leucémies (section 4.4.3).

Les résultats des micropuces pour cette partie ont été comparés à une autre étude réalisée en 2003 avec des micropuces Affymetrix par l'équipe de M.J. Weiss (Welch et al. 2004). Cette étude concluante a été effectuée en comparant le transcriptome d'une lignée cellulaire érythroblastique à différentes étapes de la différenciation. Cette lignée cellulaire, G1E, provient de cellules souches embryonnaires murines délétées pour Gata1 donc bloquées dans la différenciation érythroïde. Lors de la réexpression de Gata1 dans ces cellules, le processus de différenciation se remet à fonctionner, de façon plus homogène que par l'ajout de facteurs externes, et mime ainsi remarquablement ce qui peut se passer *in vivo* chez la souris. De plus, la micropuce Affymetrix utilisée dans l'étude de Welch et al. 2004 (Mouse Genome 430A 2.0) contient les mêmes probes que celles utilisées dans cette présente étude (Mouse Genome 430.2), ce qui permet de comparer les données. Cependant plus ancienne, Mouse Genome 430A 2.0 contient seulement 1/3 des probes présents sur la micropuce Mouse Genome 430.2, notre étude étant donc plus complète. Il existe une corrélation évidente entre les résultats de l'équipe de M.J. Weiss et ceux de notre étude: les gènes dont l'expression augmente durant la différenciation (étude de Welch et al. 2004) tendent à être sous-exprimés dans les leucémies par rapport au contrôle (notre étude), et inversement, les gènes dont l'expression diminue avec la différenciation tendent à être surexprimés par rapport au contrôle (des exemples de cette corrélation sont visibles sur les tableaux aux appendices G et H). Ceci montre que les leucémies érythroïdes induites par le rétrovirus Graffi sont dans un état plus immature que le contrôle érythroïde. Ceci est tout à fait probable. Le contrôle érythroïde provient des cellules Ter119⁺ de la moelle osseuse qui reflète tous les stades de maturation. Ainsi les oncogènes ou plus globalement les gènes impliqués dans les étapes précoces de la différenciation se trouvent certainement dans la fraction des gènes surexprimés par rapport au contrôle. Inversement, les gènes sous-exprimés par rapport au contrôle sont

certainement des gènes reliés à la fonction de la cellule mature ou aux étapes tardives de la différenciation.

4.5.2.5.1. Gènes surexprimés dans E1 E2 E3

A l'issue de l'analyse SAM, 2635 *probesets* ont passé les critères d'une valeur p inférieure à 0,01 et un FDR inférieur à 0,20 pour les gènes spécifiques de E1, E2, E3. Le tableau correspondant (appendice G) montre les résultats des 50 premiers *probesets* et est organisé de façon semblable aux tableaux précédents. Trois colonnes supplémentaires (E1-E, E2-C, E3-C) indiquent les ratios d'expression des *probesets* entre les leucémies E1, E2, E3 et le contrôle érythroïde. Une autre colonne 'G1Edb' indique les résultats de Welch et al. 2004: Une flèche montante indique que l'expression du gène a augmenté significativement au cours de la différenciation, une flèche descendante indique une diminution au cours de la différenciation, une flèche horizontale indique qu'il n'y a eu aucun changement. Dans cette colonne, les cases avec un 'moins' correspondent aux *probes* absentes dans l'analyse de Welch et al. 2004.

Les 50 premiers *probesets* correspondent à 43 gènes. Neuf gènes sont connus en relation avec la lignée érythroïde (érythrocytes, érythroblastes ou leucémies érythroïde).

Parmi ces 43 gènes, 13 sont sous-exprimés dans les leucémies par rapport au contrôle, 23 sont surexprimés, et 7 ne montrent pas de changement significatif.

4.5.2.5.1.1 Gènes sous-exprimés par rapport au contrôle

Parmi ces 13 gènes, 8 gènes sont connus dans la lignée érythrocytaire et 3 sur 13 augmentent significativement au cours de la différenciation d'après l'équipe de M.J. Weiss (G1Edb): Abcb10, Slc38a5, Ank1.

Deux grandes classes fonctionnelles divisent ses 13 gènes, l'une impliquée dans le métabolisme (transporteurs et biosynthèse du hème) et l'autre est structurelle (cytosquelette).

Gènes impliqués dans le métabolisme:

Parmi eux, Abcb10 (transporteur mitochondrial, rangs 1, 20, 28), Abcb6 (transporteur mitochondrial de porphyrine, rang 26, Krishnamurthy et al. 2006) et Cpox (rangs 10, 34, tableau 4.7) sont impliqués dans la synthèse du hème. Leur expression est connue pour augmenter avec la différenciation. Abcb10 est régulé par Gata1 (Shirihai et al. 2000). Slc38a5 (rang 11, transporteur d'acides aminés) n'est pas connu en relation avec la lignée érythrocytaire mais est augmenté significativement lors de la différenciation de G1E. Abcg4 (rang 19) est un transporteur de la famille *white*, hypothétiquement impliqué dans le transport du cholestérol. Pkd2l2 (rang 42) est un canal cationique. Plusieurs de ces transporteurs appartiennent à la sous-famille de résistance multiple aux substances exogènes (*multi-drug resistance*).

Gènes structurels:

Les gènes structurels sont Epb4.2 (rangs 4, 36, composant du cytosquelette), Tmod1 (rang 8, régulation des filaments d'actine), Ank1 (rang 24, lien entre les protéines membranaires et le cytosquelette), Spnb1 (rang 30, liaison à l'actine). Spnb1 est un gène candidat pour être impliqué dans un désordre génétique, la sphérocytose (forme anormale des globules rouges), l'anémie hémolytique la plus fréquente dans la population caucasienne (Peters et al. 2004). Epb4.2 est connu pour être régulé par le complexe multimérique Scl, E47, Gata1, Lmo2 et Ldb1 (Xu et al. 2003).

Dapk2 (rang 31) est un gène qui ne peut pas être inclus dans les 2 classes précédentes. Son rôle connu serait celui d'un suppresseur de tumeur, mais il intervient également dans la maturation des neutrophiles et il aurait un rôle dans l'érythropoïèse tardive (Wojchowski et al. 2006).

4.5.2.5.1.2 Gènes sans changement par rapport au contrôle

Ces 7 gènes sont Xtrp3s1 (rang 3, non connu), Btbd4 (rang 15, domaine en doigt de zinc), Asb-1 (rang 18, non connu), Atpv0a1 (rang 24, non connu), DNAja4 (rangs 37, 47, chaperone), Pon2 (rang 45, anti-oxydant), 1700047I17Rik (rang 50).

Aucun de ces gènes n'est encore connu en relation avec la lignée érythrocytaire, ils sont majoritairement peu connus. Un seul de ces gènes, *Pon2*, est présent dans l'analyse avec GIE et il ne montre pas non plus de changement significatif par rapport à la différenciation.

On peut penser que ces gènes qui ne semblent pas changer d'expression au cours de la différenciation ne ressortent pas dans les autres analyses de *profiling* qui comparent différents états de différenciation. L'originalité de notre étude vient du fait que nous comparons les différents types de leucémie entre eux et non uniquement par rapport au contrôle, ce qui fait ressortir la signature spécifique de chaque lignée.

Il serait intéressant d'étudier la fonction de ces gènes puisqu'elle est en général inconnue. *Asb-1* a peut-être une fonction reliée à l'hématopoïèse. *Asb1* est similaire à *Asb2* qui est régulé positivement lors de la différenciation de la lignée myéloïde HL-60 (Guibal et al. 2002).

4.5.2.5.1.2 Gènes surexprimés par rapport au contrôle

Parmi les gènes surexprimés par rapport au contrôle, *Rabgef1* (rangs 6, 25) et *Cda* (rang 7) ont été sélectionnés pour être validés en RT-PCR (description à la section 4.5.2.5.3).

Parmi les 23 gènes, peu de gènes ont des fonctions connues et 1 seul (*Mdb2*, rang 35) est connu en relation avec la lignée érythrocytaire.

Quatre gènes pourraient être liés à l'état cancéreux des cellules. Il s'agit de *Ntn4* (rang 2), *Ttc1* (rangs 17, 21), *Rps6kb2* (rang, 43), et *Pvt1* (rang 44).

Ntn4: est impliqué dans la guidance axonale et la migration cellulaire, son rôle serait aussi relié à l'angiogenèse (néo-vascularisation) (Wilson et al. 2006).

Ttc1: interagit avec Ras et active les voies de signalisation passant par la protéine Gα (Marty et al. 2003).

Rps6kb2: est activé par la voie de signalisation de l'IL3 (mTor, Pi-3Kinase) et cela confère un avantage prolifératif aux cellules (Cruz et al. 2005).

Pvt1: est impliqué dans des translocations chromosomiques de lymphomes de Burkitt et de leucémie AML. Il est situé sur le même locus que *Myc*, directement en aval et est

connu pour activer Myc chez la souris (Lennon et al. 2007). Même si ce gène est encore relativement peu étudié, il pourrait être impliqué dans la cancérisation et il serait intéressant d'étudier de façon plus approfondie son rôle oncogénique, notamment dans les leucémies érythroïdes. 4 *probesets* sont présents dans cette analyse et donnent les mêmes résultats: une très forte spécificité et un très fort ratio par rapport au contrôle.

D'après la description de leurs fonctions, 4 autres gènes peuvent susciter de l'intérêt. Il s'agit de Resp2 (rang 33), Mbd2 (rang 35), Lefty (rang 23), ou Zfp324 (rang 9).

Resp2 se lie à Ral et servirait à transmettre des signaux d'activation ainsi qu'à permettre l'endocytose de récepteur de croissance (tyrosine kinase) (Ikeda et al. 1998).

Mbd2 est impliqué dans la méthylation de l'ADN. Il permet la méthylation d'une chaîne de globine embryonnaire dans les érythroblastes pour réprimer ce gène lors du développement (Kransdorf et al. 2006).

Lefty appartient à la famille des cytokines TGF β et Zfp324 contient les domaines en doigt de zinc spécifiques des facteurs de transcription.

D'autres gènes surexprimés dans ce tableau mais non cités dans cette section sont également intéressants car leur fonction n'est pas connue.

Hemt1 (*hematopoietic cell transcript 1* ou Gml) est situé au-delà des 50 premiers *probesets* et mérite cependant d'être cité car il a une très grande spécificité érythroïde ainsi qu'un ratio -leucémies versus contrôle- des plus élevé (tableau 4.13).

Tableau 4.13: Hemt1

F1	T1	T2	T3	B1	B2	B3	M	E1	E1-E	E2	E2-E	E3	E3-C	Mk1	Mk2	Mk3
Hemt1	-1,51	-1,41	-1,17	-1,05	-1,39	-1,38	-1,44	3,11	4,02	5,05	5,96	5,64	6,56	-1,55	-1,55	-1,35

Une étude a trouvé un transcrit alternatif spécifique de la lignée érythroïde (érythroblastes) mais cette fonction n'a pas encore été étudiée en détail (Xue et al. 1999).

4.5.2.5.2. Gènes surexprimés dans E1 E2 E3 et Mk3

A l'issue de l'analyse SAM, 1795 *probesets* ont passé les critères d'une valeur *p* inférieure à 0,01 et un FDR inférieur à 0,20 pour les gènes spécifiques de E1, E2, E3 et Mk3. Le tableau correspondant (appendice H) montre les résultats des 50 premiers *probesets*. Ce tableau (appendice H) montre de bons exemples de concordance entre notre étude et les résultats de Welch *et al.* (2004). Sept gènes (*Gypa*, *Snca*, *Slc4a1*, *Alas2*, *Ferr*, *Specc1*, *Transcribed locus*) montrent une surexpression au cours de la différenciation (G1db) et sont sous-exprimés dans les leucémies par rapport au contrôle. Inversement, 3 gènes (*Gpl1a*, *Hsd3b1*, *Srxn1*) montrent une surexpression dans les leucémies par rapport au contrôle et sont réprimés lors de la différenciation (G1db).

Les gènes *Snca*, *Ltbp2*, *Gata1*, *Slamf1*, *Btbd14a* ont été choisis pour être validés par RT-PCR (section 4.5.2.5.3).

Comme précédemment, les gènes peuvent être divisés en 3 classes: gènes sous-exprimés par rapport au contrôle, gènes sans changement significatif, gènes surexprimés.

4.5.2.5.2.1 Gènes sous-exprimés par rapport au contrôle

Parmi ces gènes sous-exprimés par rapport au contrôle, 4 sont déjà connus en relation avec la lignée érythrocytaire: *Gypa* (Glycophorin A, rangs 6, 10, tableau 4.4), *Alas2* (rang 14, synthèse du hème, tableau 4.7), *Hbb-b1* (rang 15, chaîne β d'hémoglobine, tableau 4.7), *Trim10* (facteur de transcription, tableau 4.5), *Fech* (ferrochélatase, rang 42, synthèse du hème, tableau 4.7) et *Slc4a1* (rangs 11, 18).

Slc4a1 (CD233) est bien connu pour coder pour une protéine localisée à la surface des érythrocytes. Cette protéine est impliquée dans le transport du CO_2 , et elle interagit également avec GYPA. Des mutations dans le gène sont impliquées dans la sphérocytose héréditaire comme *Spnb1*.

L'homologue du gène *Specc1* (rang44), HCMOGT, fait partie d'une translocation chromosomique impliquant PDGFRB dans une leucémie myélomonocytaire juvénile (Morerio et al. 2004).

Les autres gènes n'ont pas de fonction connue et méritent d'être étudiés.

4.5.2.5.2.2 Gènes sans changement par rapport au contrôle

Parmi ces gènes, *Eraf* (rang 17), *Gata1* (rang, 15) et *Cd59a* (rangs 28, 39, section 4.5.2.2) sont connus en relation avec la lignée érythrocytaire. L'absence de changement d'expression de *Gata1* se comprend par un haut niveau d'expression dans les leucémies et les contrôles, en accord avec l'importance de ce gène dans la lignée érythrocytaire.

Eraf a comme rôle de stabiliser les chaînes libres α d'hémoglobine lors de l'érythropoïèse. Ce gène est régulé par *Gata1* et est lié à des maladies hémolytiques (β -thalassémie) (Kihm et al., 2002).

Smox (rang 13, polyamine oxidase) est une cible thérapeutique, car elle intervient dans le métabolisme des polyamines qui sont souvent impliquées dans la prolifération cellulaire et dérégulées dans les cancers. Les thérapies ciblent la suppression de la biosynthèse des polyamines tout en favorisant le catabolisme de ces molécules car cela produit des molécules de H_2O_2 cytotoxiques pour les cellules, donc capable de détruire les cellules tumorales (Wang et al. 2006).

4.5.2.5.2.3 Gènes surexprimés par rapport au contrôle

Parmi les gènes surexprimés par rapport au contrôle, *Ltbp2*, *Slamf1*, *Btbd14a* ont été sélectionnés pour être validés par RT-PCR (section 4.5.2.5.3). Six gènes codent pour des protéines avec une fonction dans la structure des cellules: 5 gènes sont impliqués dans le cytosquelette et un gène code pour une fibre de collagène. Il s'agit de:

Cdc42ep5 (rang 2): régulation de la formation des structures d'actine F

Gas2l1 (rang 4): protéine associée à l'actine

Parvb (rang 5): protéine associée à l'actine, au niveau des points d'adhésion, régule la dynamique du cytosquelette, la mobilité, la survie (Sepulveda and Wu, 2006).

Tns (rang 37): protéine associée à l'actine, points d'adhésion, migration cellulaire

Daam1 (rang 46): cytosquelette d'actine, polarité cellulaire, adhésion cellulaire, mobilité

Col5a (rang): fibre de collagène

Ainsi ces gènes impliqués dans la régulation de la forme et de la dynamique du cytosquelette complètent ceux déjà identifiés pour les gènes spécifiques de E1, E2 et E3 (section 4.5.2.5.1.1, sous-exprimés par rapport au contrôle). La forme du cytosquelette apparaît essentielle à la fonction de transporteurs de gaz des globules. Mais on peut penser également que certains des gènes identifiés ici (par exemple *Parvb* et *Daaml*) surexprimés par rapport au contrôle pourraient également refléter l'état cancéreux et permettre aux érythroblastes cancéreux de migrer et d'envahir l'organisme. Deux gènes semblent impliqués dans la signalisation, *Rap2b* (rang 43) et *Mobkl1a* (rang 50).

En conclusion sur l'analyse des gènes spécifiques des échantillons érythroïdes, 2 classes fonctionnelles sont rencontrées majoritairement lors de cette analyse: 1) fonction dans le métabolisme des globules rouges (transporteurs de métabolites, anti-oxydants et biosynthèse du hème); 2) forme et dynamique du cytosquelette. Cela reflète l'unicité fonctionnelle des globules rouges dans le transport des gaz de l'organisme. Cette observation semble plus homogène et plus simple comparée aux fonctions mégacaryocytaires et plaquettaires. La fonction du globule rouge se fait par une bonne fixation des molécules de gaz sur l'hémoglobine. La forme des globules rouge est importante pour cette fixation ainsi que pour le transport rapide dans les vaisseaux sanguins et capillaires fins des bronches. Le globule rouge doit lutter contre différents stress oxydatifs. La régulation de sa fonction se fait principalement par régulation du nombre de globules rouges et donc par une augmentation de l'érythropoïèse (stimulation par l'érythropoïétine).

Les leucémies érythroïdes sont associées à une mauvaise réponse aux traitements thérapeutiques. Se peut-il que la présence de nombreux transporteurs ABC de classe MDR/TAP puissent jouer un rôle dans cette faible efficacité des traitements? (voir également implication de *Cda*, section 4.5.2.5.3).

Par rapport aux leucémies mégacaryocytaires, moins d'oncogènes ont été révélés dans l'analyse. Ainsi à ce moment de l'analyse, les gènes majeurs connus apparemment impliqués dans les leucémies érythroïdes peuvent être *Gata1*, *Fog1*, *Fli1*, *NF-E2*, *Scl*, *Lmo2* (tableau 4.5) qui sont spécifiques des leucémies érythroïdes et mégacaryoblastiques. *Fli-1* et *Nfe2* sont des cibles d'intégrations du rétrovirus de Friend et sont les premiers intervenants dans le déclenchement des leucémies érythroïdes induites par ce rétrovirus. Quelques échantillons de nos tumeurs ont été testés pour le réarrangement de *Fli-1* dans les leucémies érythroïdes induites par le rétrovirus Graffi et un réarrangement a été

identifié dans 50% des cas (discussion chapitre II). Il serait intéressant de tester le réarrangement de *Fli-1* dans *E1*, car elle est la seule à surexprimer ce gène (tableau 4.5).

Il ressort de cette analyse que beaucoup de gènes ont une fonction associée inconnue et que qu'il serait donc intéressant de pouvoir élucider leur fonction.

4.5.2.5.3. Gènes sélectionnés pour les tests de RT-PCR

De même que pour les leucémies mégacaryoblastiques, des gènes spécifiques des leucémies érythroïdes ont été sélectionnés pour être testés en RT-PCR. Il s'agit de *Snca*, *Ltbp2*, *Gata1*, *Btbd14a*, *Cda* et *Rabgef1*. Le tableau 4.14 ci-dessous présente les résultats des micropuces (SAM et ratios) pour les différents *probesets* associés à ces gènes.

Tableau 4.14: Gènes érythrocytaires sélectionnés pour l'analyse par RT-PCR

tableau	nom	probesets	SAM d-score	P-value	FDR	rangs	T1	T2	T3	B1	B2	B3	M	E1	E2	E3	Mk1	Mk2	Mk3
4.20	<i>Snca</i>	1	10,2	<0,001	0,02	19	-3,3	-0,2	-1,0	-2,5	-3,2	-1,8	-1,5	4,6	4,2	1,7	-1,0	-0,7	2,7
4.20	<i>Snca</i>	2	10,2	<0,001	0,02	19	-3,3	-0,2	-1,0	-2,5	-3,2	-1,8	-1,5	4,6	4,2	1,7	-1,0	-0,7	2,7
4.20	<i>Snca</i>	3	3,5	0,007	0,13	1505	-0,3	-0,6	-0,5	-0,3	-0,3	-0,4	-0,4	1,7	0,6	1,7	0,0	-0,3	0,3
4.20	<i>Snca</i>	4	3,3	0,009	0,14	1730	-0,6	-0,7	-0,5	-0,7	-0,8	-0,5	-0,8	3,2	0,9	1,7	-0,3	-0,7	0,7
4.20	<i>Ltbp2</i>	1	15	<0,001	0,01	3	-0,7	-0,7	-0,7	-0,7	-0,8	-0,7	-0,9	2,2	1,5	1,7	-0,5	-0,8	1,4
4.20	<i>Gata1</i>	1	10,6	<0,001	0,02	16	-1,9	-1,9	-2,1	-2,1	-2,3	-2,7	-1,9	4,5	4,1	1,7	0,6	-1,4	2,5
4.20	<i>Btbd14a</i>	1	9,5	<0,001	0,02	25	-0,6	-0,9	-0,9	-0,9	-0,8	-0,9	-0,5	2,3	1,5	1,7	-0,4	-0,4	1,0
4.20	<i>Btbd14a</i>	2	8,3	<0,001	0,02	52	-1,0	-1,3	-1,1	-1,1	-1,2	-0,8	-0,4	3,0	1,4	1,7	-0,8	-0,7	1,7
4.20	<i>Btbd14a</i>	3	7,8	<0,001	0,03	71	-1,1	-1,1	-1,1	-1,0	-1,3	-1,2	-0,4	2,8	1,7	1,7	-0,2	-0,4	1,2
4.20	<i>Btbd14a</i>	4	3,6	0,005	0,12	1354	-0,4	-0,2	0,1	-0,2	-0,1	-0,3	0,0	0,1	0,2	1,7	0,1	0,1	0,3
4.20	<i>Slamf1</i>	1	9,9	<0,001	0,02	21	-1,3	-1,6	-1,7	-1,2	-1,6	-1,5	-2,0	4,1	3,0	1,7	-0,3	-1,2	1,9
4.20	<i>Slamf1</i>	2	9,5	<0,001	0,02	24	-1,0	-2,6	-1,9	-0,8	-1,2	-0,9	-2,7	3,9	2,9	1,7	-0,3	-0,7	2,2
4.20	<i>Slamf1</i>	3	8,9	<0,001	0,02	40	-0,9	-2,8	-2,2	-0,9	-1,2	-1,4	-3,3	4,5	3,1	1,7	-0,4	-0,7	2,6
4.18	<i>Cda</i>	1	25,0	<0,001	0,01	7	-1,5	-1,6	-1,6	-1,7	-1,8	-0,6	-1,3	4,3	4,8	4,9	-1,6	-1,7	-0,7
4.18	<i>Rabgef1</i>	1	25,0	<0,001	0,01	6	-0,5	-0,8	-0,4	-0,9	-1,3	-1,2	-0,8	2,2	2,3	2,3	-0,4	-0,4	-0,1
4.18	<i>Rabgef1</i>	2	18,1	<0,001	0,01	25	-1,2	-0,8	-0,5	-0,9	-0,7	-1,0	-0,7	1,9	2,1	2,2	-0,5	-0,3	0,2
4.18	<i>Rabgef1</i>	3	15,7	<0,001	0,01	57	-0,9	-1,1	-0,5	-1,0	-0,8	-1,3	-0,5	2,1	2,5	2,5	-1,3	0,0	0,2

4.5.2.5.3.1 *Snca* (Synuclein, alpha)

Snca est localisé sur le chromosome 6 chez la souris. Les protéines murines et humaines partagent 95% d'homologie. *Snca* est très connu pour son implication dans la maladie de Parkinson: une association entre des mutations de *Snca* et cette maladie a été établie (Bonini et al. 2005). De plus, la protéine codée par *Snca* s'accumule de façon anormale dans les corps de Lewy dans le cerveau des patients atteints de cette maladie (Bonini et al. 2005). La localisation de *Snca* dans le cerveau est pré-synaptique et une étude montre une association de *Snca* avec une molécule chaperonne qui jouerait un rôle dans le repliement des protéines SNARE qui régulent le trafic vésiculaire au niveau de la

synapse (Bonini et al. 2005). Cependant, la fonction normale de Snca dans le cerveau ne semble pas encore élucidée. Très récemment, en juin 2007, une étude décrit pour la première fois l'expression de Snca dans les érythrocytes et érythroblastes (Nakai et al. 2007). Cependant, sa fonction dans la lignée érythroïde est encore complètement inconnue. Les résultats des micropuces confirment l'étude de Nakai et al., et le modèle Graffi peut donc être utilisé pour tenter d'élucider la fonction de la protéine codée par Snca.

4.5.2.5.3.2.Ltbp2 (latent transforming growth factor beta binding protein 2)

Ltbp2 est situé sur le chromosome 12 chez la souris et les protéines murine et humaine partagent 80% d'homologie. Il existe 4 protéines LTBP, nommées de LTBP1 à LTBP4. Les protéines LTBP se complexent avec les molécules sécrétées de TGF β afin de l'inactiver jusqu'au lieu où TGF β doit exercer son action (Oklu et al. 2000). TGF β est une cytokine qui a de nombreux effets, notamment capable d'inhiber ou d'activer la prolifération cellulaire. Il était alors important de connaître l'expression des différents isoformes de Tgf β et de Ltbp dans les leucémies érythroïdes induites par le rétrovirus murin Graffi (tableau 4.15).

Tableau 4.15: Expression des Ltbps et Tgf β

	T1	T2	T3	B1	B2	B3	M	E1	E2	E3	Mk1	Mk2	Mk3
Tgfb1	-2,19	-2,97	-0,76	0,24	-0,28	0,01	0,01	1,80	1,48	1,58	0,34	-0,43	1,18
Tgfb2	-0,02	-0,24	0,02	0,09	-0,04	-0,14	0,06	0,23	0,24	-0,04	0,05	-0,11	-0,09
Tgfb3	0,02	-0,09	0,09	0,05	-0,07	0,05	-0,21	-0,14	0,09	-0,28	0,12	0,43	-0,06
Ltbp1	-0,70	-0,68	-0,32	-0,42	-0,71	-0,49	-0,67	3,18	-0,55	0,28	0,08	0,21	0,78
Ltbp1	-1,74	-1,56	-1,73	-1,69	-1,64	-1,82	-1,45	4,45	-0,87	3,13	-0,34	2,87	2,41
Ltbp2	-0,75	-0,74	-0,71	-0,72	-0,84	-0,74	-0,86	2,15	1,49	1,68	-0,53	-0,83	1,40
Ltbp3	-0,44	-0,58	0,02	-0,40	-0,55	-0,50	-0,78	-0,50	-0,01	-0,25	-0,01	2,89	1,09
Ltbp4	-0,18	-0,09	-0,10	0,03	0,42	0,02	-0,18	-0,20	0,03	0,07	0,27	-0,05	-0,04

D'après le tableau 4.15 ci-dessus, l'isoforme Tgfb1 est celui qui est exprimé spécifiquement dans les 3 leucémies érythroïdes et Mk3, de façon identique à Ltbp2. On peut donc émettre l'hypothèse que Ltbp2 interagit directement avec Tgfb1 dans les leucémies érythroïdes induites par le rétrovirus Graffi. Certaines études ont effectivement montré une co-ordination dans l'expression d'un isoforme Ltbp et d'un isoforme Tgf β précis, notamment de Ltbp1 et Tgfb1 dans les cellules HEL (Oklu et al. 2000) mais aucune étude n'a encore montré d'interaction de Ltbp2 avec un isoforme de Tgf β . Les

protéines LTBP auraient un rôle indépendant de Tgfb (Oklu et al. 2000). Ainsi la protéine LTBP2 pourrait jouer le rôle d'une protéine fibrillaire impliquée dans l'adhésion cellulaire (Vehvilainen et al. 2003).

4.5.2.5.3.3. Btbd14a (BTB/POZ domain containing 14A)

Le gène est situé sur le chromosome 2 chez la souris et les protéines humaine et murine partagent 56% d'homologie. De nombreux E.S.Ts correspondent à ce gène mais *sa fonction n'a jamais été étudiée chez la souris*. En revanche, son homologue humain BTBD14A (NAC1) est plus connu et semble avoir plusieurs propriétés d'un oncogène. BTBD14A a été découvert pour la première fois dans le cerveau de rats auxquels on avait administré de la cocaïne (Korutla et al. 2002). Son rôle potentiellement oncogénique a d'abord été suggéré par la présence du domaine BTB/POZ qui est commun à de nombreux oncogènes impliqués notamment dans les néoplasies hématopoïétiques comme Bcl6 par exemple (Nakayama et al. 2006).

Des études ont alors révélé la surexpression de BTBD14A dans différents types de cancer, comme ceux de l'ovaire et du sein. Globalement, il a été trouvé que BTBD14A influence la prolifération cellulaire, la survie des cellules et il semblerait également impliqué dans l'induction des tumeurs dans un modèle de cancer de l'ovaire et dans des souris *nude* (Nakayama et al. 2006). Il n'a pas encore été observé dans les leucémies érythroïdes, murine ou humaine. Il est donc intéressant de valider son profil érythroïde par RT-PCR et de poursuivre son étude fonctionnelle dans ce type spécifique de leucémie.

4.5.2.5.3.4. Slamf1 (signaling lymphocytic activation molecule family member 1)

Le gène est situé sur le chromosome 1 chez la souris et les protéines humaine et murine partagent 60% d'homologie. Slamf1 (Cd150) code pour un récepteur membranaire de la superfamille des immunoglobulines. Il est classiquement associé aux cellules T et B activées, mais une étude approfondie chez l'Humain montre qu'il est aussi présent sur les cellules T et B non activées, les monocytes et les cellules dendritiques (Romero et al. 2004). Par contre, il n'a pas été détecté sur les érythrocytes humains non leucémiques (Romero et al. 2004). Les résultats des micropuces montrent une présence de Slamf1 dans

tous les échantillons (résultats MAS5.0 non montrés) mais une très forte surexpression de *Slamf1* dans les leucémies érythroïdes induites par le rétrovirus Graffi et ceci pour les 3 *probesets* disponibles pour ce gène. Il est donc très peu probable que cela soit un faux-positif. Il serait certainement intéressant de voir si *Slamf1* est également surexprimé dans d'autres modèles de leucémies érythroïdes. Une étude récente a montré pour la première fois que CD150 était aussi exprimée sur les cellules souches hématopoïétiques et non sur les cellules progénitrices pluripotentes et que cette expression pourrait permettre de bien séparer et trier ces 2 populations (Kiel et al. 2005).

Les voies de signalisation et donc les effets passant par l'activation du récepteur membranaire SLAMF1, dans les cellules leucémiques induites par le rétrovirus Graffi, sont sans aucun doute très différents de ce qui se passe au niveau lymphocytaire. En effet, des études montrent par exemple que les voies de signalisation et les effets de *Slamf1* sont différents dans les cellules lymphoïdes par rapport aux cellules dendritiques ou monocytaires. Ainsi, dans les lymphocytes T, SLAMF1 agit avec une protéine nommée SAP codée par le gène *Sh2d1a* (Romero et al. 2004). D'après le résultat des micropuces, *Sh2d1a* est effectivement surexprimé spécifiquement et uniquement dans les leucémies de type T (tableau 4.16).

Tableau 4.16: Expression de *Sh2d1a*

	T1	T2	T3	B1	B2	B3	M	E1	E2	E3	Mk1	Mk2	Mk2
<i>Sh2d1a</i>	5,4	5,8	5,0	-1,5	-1,6	-1,7	-1,5	-1,7	-1,6	-1,7	-1,5	-1,6	-1,8

Dans les cellules présentatrices d'antigènes (monocytes et cellules dendritiques), SLAMF1 semblerait interagir avec une autre protéine EAT-2 (*Sh2db1,2*) (Romero et al.2004). Aucun *probeset* Affymetrix ne correspondant à ce gène, il n'est donc pas possible de vérifier une potentielle expression de ce gène dans les leucémies érythroïdes.

Il serait important de connaître quels avantages donnent la surexpression de *Slamf1* aux cellules leucémiques érythroïdes ainsi que d'établir les voies de signalisation reliées à ce récepteur.

4.5.2.5.3.5. Cda (Cytidine deaminase)

Le gène est situé sur le chromosome 4 chez la souris et les protéines murine et humaine partagent 81% d'homologie. La protéine codée par ce gène est responsable de la résistance de certaines cellules aux traitements de chimiothérapie par des composés comme la cytarabine, la gencitabine ou ara-C (Ge et al. 2005). Ces agents thérapeutiques sont normalement transformés en métabolites actifs (analogues de nucléosides) dans la cellule et vont bloquer la synthèse d'ADN. La CDA va les transformer en métabolites inactifs qui ne peuvent plus bloquer la synthèse d'ADN.

Il est particulièrement important de s'intéresser à l'expression de ce gène dans les leucémies induites par le rétrovirus Graffi. D'une part, Cda est fortement exprimé dans les leucémies érythroïdes induites par le rétrovirus Graffi. D'autre part, des études montrent une corrélation entre les leucémies mégacaryoblastiques (AMKL) des patients atteints du syndrome de Down et une plus faible expression de Cda. En effet, il est connu depuis de nombreuses années que les patients atteints du syndrome de Down et d'une leucémie mégacaryoblastique répondaient mieux au traitement alors que les patients non atteints du syndrome de Down et souffrant du même type de leucémie avaient de très faibles chances de guérison (Hitzler 2005). Cela semble être corrélé à une moins forte expression de Cda et une moins forte activité de la protéine dans les leucémies AMKL des patients souffrant du syndrome de Down (Ge et al. 2005). Gata1 est muté uniquement dans les leucémies AMKL des patients souffrant du syndrome de Down (Hitzler 2005). Une étude suggère alors également fortement une corrélation entre Gata1 et Cda: Gata1 sauvage stimulerait l'expression de Cda (Ge et al. 2005). Ainsi il est important d'observer précisément l'expression de Cda dans les leucémies Mk induites par le rétrovirus Graffi et de vérifier également s'il existe une corrélation entre l'expression de Gata1 et Cda globalement dans tous les échantillons leucémiques. D'après les résultats des micropuces, l'expression de Cda ne semble pas corrélée à celle de Gata1 mais une validation RT-PCR sur d'autres échantillons est nécessaire. Il serait très intéressant de connaître les facteurs de transcription qui régulent l'expression de Cda car cela peut être très important pour améliorer la sensibilité aux thérapies des leucémies érythroïdes et mégacaryoblastiques humaines qui sont associées à de très faibles chances de guérison.

4.5.2.5.3.6. Rabgef1 (RAB guanine nucleotide exchange factor (GEF) 1)

Le gène est situé sur le chromosome 5 chez la souris et les protéines souris et humaines partagent 95% d'homologie. Les protéines 'GEF' (*guanine nucleotide exchange factor*) ont comme rôle d'activer les protéines de la famille RAS. Les protéines RAB font partie de la famille RAS et elles sont impliquées dans le contrôle du trafic vésiculaire à l'intérieur des cellules. Rabgef1 (Rabex-5) interagit avec les protéines de la sous-famille RAB5 qui incluent RAB5, RAB21 et RAB22 (Mattera et al. 2006). La sous-famille Rab5 semble impliquée dans la fusion des vésicules avec les endosomes.

Le résultat des micropuces montrent que, parmi les gènes de la sous-famille Rab5, Rab22a est celui qui est spécifiquement exprimé dans les leucémies érythroïdes induites par le rétrovirus Graffi (tableau 4.17).

Tableau 4.17: Expression de Rabgef1 et Rab22a

	T1	T2	T3	B1	B2	B3	M	E1	E2	E3	Mk1	Mk2	Mk3
Rabgef1	-0,9	-1,1	-0,5	-1,0	-0,8	-1,3	-0,5	2,1	2,5	2,5	-1,3	0,0	0,2
Rabgef1	-1,2	-0,8	-0,5	-0,9	-0,7	-1,0	-0,7	1,9	2,1	2,2	-0,5	-0,3	0,2
Rabgef1	-0,5	-0,8	-0,4	-0,9	-1,3	-1,2	-0,8	2,2	2,3	2,3	-0,4	-0,4	-0,1
Rab22a	0,5	-0,3	-0,5	-0,2	-0,5	-0,3	-0,4	1,1	1,0	1,0	-0,5	-0,3	-0,6
Rab22a	0,0	-0,3	-0,3	-0,3	-0,5	-0,2	-0,3	1,0	0,9	1,2	-0,7	-0,4	-0,3
Rab22a	0,1	-0,3	-0,2	-0,2	-0,6	-0,5	-0,2	0,9	1,5	1,4	-0,9	-0,9	-0,1

On peut donc émettre l'hypothèse que Rabgef1 active spécifiquement Rab22a dans les leucémies érythroïdes induites par le rétrovirus Graffi. Rab22 localise avec les endosomes ainsi qu'avec le trans-Golgi et son rôle probable serait de réguler le transport entre le trans-Golgi et les endosomes. Il est connu que Rabgef1 est exprimé dans les cellules mastocytaires où il joue sur la dégranulation. Il semble être régulé par ubiquitination (Mattera et al. 2006). Ni Rabgef1 ni Rab22a n'ont encore été rapportés spécifiques de la lignée érythroïde et/ou des leucémies érythroïdes et il serait donc intéressant de poursuivre l'étude fonctionnelle des protéines codées par ces 2 gènes.

4.5.2.6. Gènes spécifiques des leucémies érythroïdes et mégacaryocytaires

La dernière analyse concernant les leucémies non-lymphoïdes présentée dans ce chapitre est celle correspondant aux gènes spécifiques des échantillons érythroïdes et mégacaryoblastiques. 2077 *probesets* ont passé les critères de l'analyse SAM et le tableau correspondant (appendice I) en présente les 50 premiers. Les 50 premiers *probesets* codent 32 gènes.

Deux gènes (Trfr2, Tal1) sont déjà connus en relation avec la lignée érythro- et mégacaryocytaire. Tal1 (rang 6) n'est autre que Scl (tableau 4.5). Trf2 (rangs 1 et 2) ressemble à Trf1, récepteur 1 de la transferrine (CD71), impliqué dans le transport du fer à l'intérieur de la cellule. La distribution tissulaire de TFR2 est différente de celle de TRF1 suggérant un rôle unique. L'expression de Trf2 diminue durant la différenciation érythroïde (Trinder et al. 2003), ce qui est en accord avec le résultat des micropuces avec une surexpression de ce gène par rapport au contrôle érythroïde. L'expression a aussi été détectée sur les plaquettes humaines bien que son rôle n'ait pas été étudié (Hannuksela et al. 2003).

Sept gènes sont associés à la lignée érythrocytaire:

Sox6 (rangs 5, 14, 26, 46): facteur de transcription impliqué dans l'érythropoïèse (survie des cellules érythroïdes, prolifération, différenciation terminale, régulation du gène de la globine, Cohen-Barak et al. 2007).

Spn1 (rangs 10, 11): spectrine, composant majeur du squelette des globules rouges. Des mutations du gène peuvent conduire à la sphérocytose.

Aq9 (rangs 11, 43): augmente avec la différenciation érythroïde, permet l'ingestion de glycérol par les globules rouges (Liu et al. 2007).

Rhag (rang 12): protéine membranaire à la surface des globules rouges (Rhésus).

Podxl (rang 13): sa fonction n'est pas très connue, l'érythropoïétine peut induire son expression dans les cellules de la moelle osseuse (Sathyanarayana et al. 2007). Son expression est aussi corrélée à l'agressivité de certains cancers.

Nt5c3 (rang 17): sa présence dans les érythrocytes est connue. Une mutation de ce gène peut entraîner une anémie hémolytique.

Dyrk3 (rang 28): sa présence dans les érythrocytes est connue mais sa fonction n'a pas été étudiée (Zhang et al. 2005).

Trois gènes sont connus dans la lignée mégacaryocytaire.

Rab38 (rangs 19, 48): est connu pour réguler les granules denses plaquettaires et il est aussi associé au syndrome hémorragique de Hermanski-Pudlak (Di Pietro and Dell Angelica, 2005).

Gp5 (rang 16): CD42d forme le facteur de Willebrand avec d'autres glycoprotéines plaquettaires (récepteur au fibrinogène).

Rab27b (rang 27): a un rôle dans la sécrétion des plaquettes (Tolmachova et al. 2007).

Certains autres gènes de ce tableau n'ont pas de fonction connue et il serait pourtant intéressant de les étudier.

Quatre gènes qui ne sont pas reliés à la lignée érythrocytaire ou mégacaryocytaire méritent d'être étudiés. Il s'agit de Ptpn12 (protéine tyrosine phosphatase, mobilité, oncogénèse), Alox5 (augmenté dans les cancers, synthèse des leucotriènes), Ptger (récepteur couplé aux protéines G) et Aldh1a1. Ce dernier est particulièrement intéressant. Il se trouve au rang 3 du tableau. Il est surexprimé avec la même intensité dans les échantillons E et Mk et sous-exprimé de façon égale dans les autres échantillons. Ce gène code pour la seconde enzyme de l'oxydation de l'alcool et permet la conversion du rétinol en acide rétinoïque (agissant sur la protéine RAR α). Il a été récemment identifié que ce gène est exprimé par les cellules souches hématopoïétiques et il est maintenant un marqueur de ces cellules souches. La protéine codée est reconnue maintenant comme un régulateur clé de la différenciation des cellules souches (Chute et al. 2006).

Au-delà des 50 premiers probesets, l'oncogène Pim-1 apparaît spécifique des leucémies érythroïdes et mégacaryoblastiques et le ratio leucémies érythroïdes versus contrôle ne montre pas de changement significatif (non montré). Pim-1 est une sérine / thréonine kinase cytoplasmique souvent dérégulée dans les leucémies. Son rôle normal dans l'érythropoïèse (Chiocchetti et al. 2005) et la mégacaryopoïèse (Kim et al. 2002) reste flou.

Les gènes présentés dans ce tableau représentent les gènes communs aux leucémies érythroïdes et mégacaryoblastiques *induites par le rétrovirus Graffi*. Il est impossible de conclure que ces gènes sont normalement exprimés dans les lignées, érythroïde ou

mégacaryocytaire, non leucémiques. Ce tableau traduit certainement la promiscuité existante entre la lignée érythrocytaire et mégacaryocytaire qui provoque l'apparition de caractères biphénotypiques lors de la cancérisation à partir d'un progéniteur érythroïde, mégacaryocytaire ou bipotent. Plusieurs exemples (notamment E1 et Mk3, et chapitre II) ont permis de voir le caractère souvent biphénotypique des leucémies érythroïdes et mégacaryocytaires induites par le rétrovirus Graffi. Ce phénomène cependant semble aussi observé chez l'humain, les lignées cellulaires obtenues à partir de ces 2 types de leucémies (AM6 ou AML7) montrent souvent des caractères biphénotypiques. Les gènes identifiés dans ce tableau n'en sont pas moins intéressants à étudier.

4.5.3. Analyse ciblée sur les leucémies lymphoïdes

Cette section présente l'analyse des résultats des micropuces ciblée sur l'analyse des leucémies lymphoïdes. Les résultats pour les leucémies T sont présentés en premier, suivis de l'analyse des résultats pour les leucémies B, des résultats pour les gènes communs des leucémies T et B et quelques exemples de gènes sélectionnés pour être validés par RT-PCR.

4.5.3.1 Les leucémies T

4.5.3.1.1. Rappel des caractéristiques des leucémies T

Les leucémies T induites par le rétrovirus Graffi sont très hétérogènes. Suivant les marqueurs phénotypiques utilisés, 3 groupes se distinguent: $CD4^+CD8^+$, $CD4^+CD8^-$ et $CD4^-CD8^+$. Les phénotypes $CD4^+CD8^+$ et $CD4^+CD8^-$ sont les plus fréquents. L'hétérogénéité ne repose pas seulement sur les marqueurs CD4 et CD8. Par exemple, on observe des leucémies réarrangées pour les récepteurs des cellules T et des cellules B et d'autres non réarrangées (chapitre II).

Un autre élément caractérisant les leucémies T induites par le rétrovirus Graffi est que 2 populations co-existent souvent dans l'organisme des souris souffrant de ce type de leucémies: une population $CD4^+CD8^+$ et $CD4^+CD8^-$, ou une population $CD4^+CD8^+$ et $CD4^-CD8^+$. On ne retrouve jamais la co-existence des phénotypes $CD4^+CD8^-$ et $CD4^-$

CD8⁺. On peut penser que dans les 2 populations co-existantes, le phénotype CD4⁺CD8⁺ est le plus immature.

Afin de réaliser les micropuces, une leucémie de chaque type a été choisie: T1 (CD4⁺CD8⁺), T2 (CD4⁺CD8⁺), T3 (CD4⁺CD8⁺). Chacune des populations leucémiques était unique dans l'organisme des souris dont elles provenaient. Ces 3 phénotypes ont été choisis intentionnellement parce qu'ils sont distincts. Ainsi cela augmente la potentialité que les gènes communs entre les 3 types de leucémies T représentent de façon plus absolue ce type de leucémie, et cela permet également de mieux trouver les gènes spécifiques des autres types de leucémies.

Le contrôle (CT) est constitué de cellules du thymus triées avec CD4⁺ et CD8⁺. Il s'agit très majoritairement de cellules CD4⁺CD8⁺ (97%) avec une petite population de CD4⁺CD8⁻ et CD4⁻CD8⁺. On peut s'attendre à ce que le contrôle ait plus de similitudes avec la leucémie T1 portant le même phénotype.

4.5.3.1.2. Vérification des phénotypes de T1, T2, T3

Le tableau ci-dessous (tableau 4.18) présente l'expression des marqueurs CD3, CD4 et CD8 par les échantillons leucémiques.

Tableau 4.18: Phénotypes des leucémies T														
	CD4+CD8+	CD4-CD8+	CD4+CD8-	CD4+CD8+										
	T1	T2	T3	CT	B1	B2	B3	M	E1	E2	E3	Mk1	Mk2	Mk3
Cd4	1,54	-0,98	3,37	4,99	-0,98	-0,98	-1,01	-0,87	-0,69	-0,72	-0,95	-0,88	-0,93	-0,63
Cd4	0,72	-0,53	3,03	4,38	-0,75	-0,80	-0,86	-0,75	-0,62	-0,80	-0,57	-0,86	-0,83	-0,57
Cd8a	3,28	5,45	-1,30	6,43	-1,43	-1,47	-0,99	-0,97	-1,26	-1,39	-1,13	-1,02	-1,33	-1,31
Cd8a	4,00	5,96	-1,49	7,49	-1,02	-1,12	-1,54	-1,43	-1,41	-1,57	-1,58	-1,05	-1,50	-1,21
CD3e	4,01	4,10	4,30	5,88	-1,76	-1,56	-1,18	-1,64	-1,48	-1,80	-1,75	-0,92	-1,79	-1,34
CD3d	5,39	5,39	5,25	5,68	-2,91	-2,69	-2,81	-0,65	-2,11	-2,47	-2,57	-1,29	-2,55	-1,50

Comme on peut le voir sur le tableau, le gène Cd4 est bien surexprimé de façon spécifique par T1, T3 et CT. Cd8 est exprimé spécifiquement par T1, T2 et CT. On peut voir aussi que les thymocytes (CT) montrent une expression plus forte des marqueurs par rapport aux échantillons leucémiques.

Les analyses des micropuces permettent également d'identifier d'autres marqueurs spécifiques (CD: Cluster of Differentiation) sur-exprimés dans T1, T2 et T3. Il s'agit de Cd6, Cd2, Cd5, Cd160, Cd28 et Cd96 (appendice A). Ils sont tous connus pour être exprimés à la surface des lymphocytes T (Wang et al. 1992, Starling et al. 1996, Wakkach et al. 2001, Bour-Jordan et Blueston 2002, Raman 2002, Tsujimura et al. 2006). Parmi ces différents marqueurs, Cd6 et Cd28 sont les plus spécifiques des leucémies T induites par le rétrovirus Graffi.

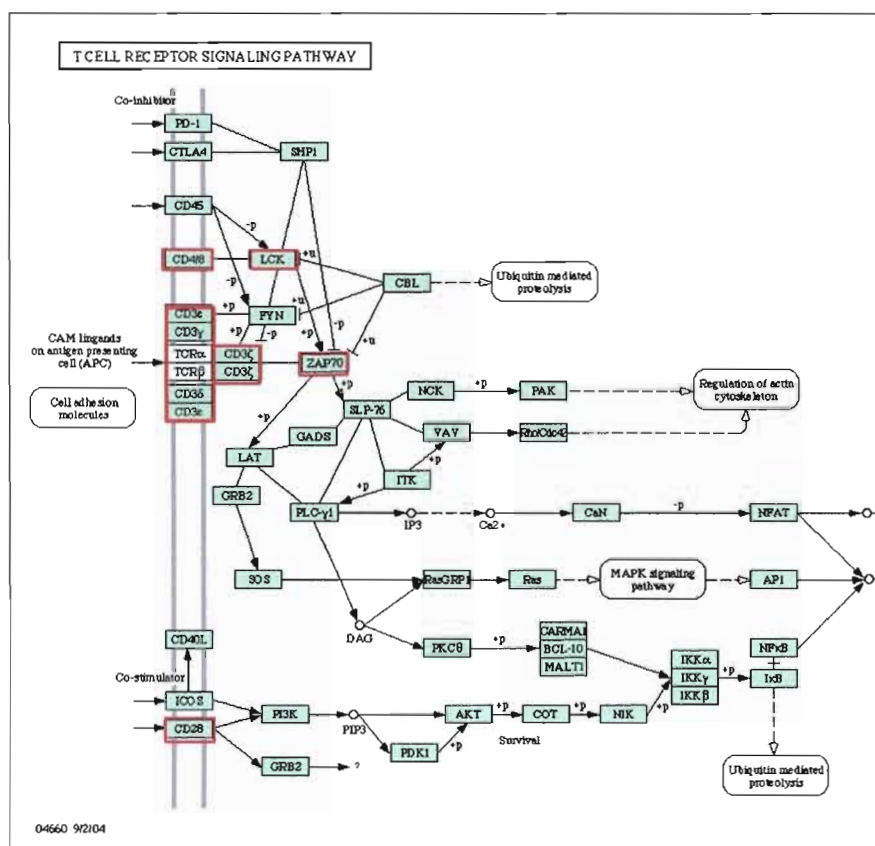
Cd160, surexprimé dans les 3 leucémies T, est intéressant car il est normalement peu exprimé par les cellules T normales. En effet, Cd160 est seulement exprimé par les cellules T $\gamma\delta$ et une sous-classe de cellules $CD4^+CD8^+$ chez l'humain et la souris (Tsujimura et al. 2006). Effectivement, il n'est pas exprimé par les thymocytes double-positifs (contrôle CT). En revanche, il est exprimé dans les 3 types de leucémies. Il serait intéressant d'étudier son expression dans un échantillonnage plus grand de leucémies T et notamment de leucémies T humaines.

4.5.3.1.3. Gènes surexprimés dans les leucémies T

A l'issue de l'analyse SAM, 1472 *probesets* ont passé les critères d'une valeur p inférieure à 0,01 et une valeur de FDR inférieure à 0,20.

Le tableau 27 (p. 302-3) montre les 50 premiers *probesets*. Ils correspondent à 35 gènes. Parmi ces 35 gènes, 22 codent pour des protéines dont le rôle dans les lymphocytes T est connu. Il s'agit majoritairement pour ces 35 gènes de récepteurs membranaires ou de protéines impliquées dans la transmission du signal. Plus précisément, on retrouve tous les composants du complexe TCR/CD3 qui reconnaît les antigènes et transmet le signal d'activation à l'intérieur de la cellule (chaînes α et β du TCR et chaînes β , δ , γ , ϵ , ζ de CD3). La voie de signalisation du TCR est illustrée sur la figure 4.12 et les gènes cités ci-dessus sont encadrés en rouge. Zap70 et Lck codent aussi pour 2 protéines très importantes pour la transmission de signaux des cellules T (figure 4.12). Elles se retrouvent au rang 460 pour Lck, et 129 et 1470 pour Zap70 des gènes spécifiques des leucémies T (non montré). Ceci signifie que des résultats significatifs peuvent être trouvés même dans les rangs inférieurs de la liste issue de l'analyse SAM et donc que le critère de sélection prenant une valeur p inférieure à 0,01 semble être adéquat.

Tcf7, Tcrim, Gata3, Itk, que l'on retrouve dans les rangs 1, 3, 10 et 11 respectivement de l'analyse SAM, ont également un rôle relié à l'activation du TCR (Oosterwegel et al. 1991, Kirchgessner et al. 2001, Grasis et al. 2003, Hernandez-Hoyos et al. 2003).



(schéma tiré de la banque de données Kegg, <http://www.genome.ad.j>)

Figure 4.12: Gènes impliqués dans la voie de signalisation du TCR (récepteur des cellules T). Les gènes encadrés en rouges sont spécifiquement surexprimés dans les leucémies T induites par le rétrovirus Graffi.

Une autre donnée importante est la comparaison de chaque leucémie T avec le contrôle (CT) (appendice J, colonnes 'T1-C', 'T2-C', 'T3-C'). Dans les 50 premiers *probesets*, 7 gènes sont surexprimés par rapport au contrôle dans les 3 conditions (Tcrγ-7, 1700017B05Rik, Glak1, 2010100012Rik, Pomc1, Mphos9, 1500009L16Rik) et 7 gènes sont sous-exprimés dans toutes les conditions (Tcf7, CD3z, CamK4, Thy1, E430004N04Rik,

Cd6, Tnfrsf7). Les autres gènes ne présentent pas de variations significatives par rapport au contrôle.

Le contrôle T (les thymocytes) sont dans un état immature. Le fait que l'expression des gènes caractéristiques de la lignée T (cités plus haut), soient sous-exprimés dans les leucémies par rapport au contrôle laisse supposer que les leucémies sont tout de même dans un état plus immature que le contrôle. Le contrôle idéal devrait en fait être constitué de 2 échantillons, un échantillon de cellules contrôles immatures et d'un échantillon de cellules contrôles matures afin de bien interpréter les résultats des ratios leucémie versus contrôle. On peut supposer que les oncogènes se trouvent majoritairement dans les gènes surexprimés dans les leucémies par rapport au contrôle. Cependant, on ne peut négliger que certains gènes qui bloquent la différenciation et qui sont normalement exprimés dans les cellules immatures se trouvent certainement dans la fraction inchangée du ratio leucémie versus contrôle

Parmi les gènes spécifiques des leucémies T et surexprimés par rapport au contrôle, voici l'exemple de 5 gènes qui suscitent de l'intérêt et dont l'étude de fonction pourrait certainement apporter de nouvelles connaissances concernant la lignée T ou le processus de leucémogenèse de la lignée T. Ces gènes sélectionnés sont situés majoritairement au-delà des 50 premiers *probesets*. En effet, au-delà des 50 premiers *probesets*, les résultats des ratios sont encore très convaincants et beaucoup de gènes ne sont pas connus en relation avec la lignée T.

Mphos9 (M-phase phosphoprotein 9): semble impliqué dans le cycle cellulaire (Matsumoto-Taniura et al. 1996). 4 *probesets* sur 4 sont présents dans le tableau (rangs 47, 50, 69, 82) et peu de choses sont connues globalement sur ce gène.

Josd3 (Josephin domain contain 3): 7 *probesets* sont présents dans la liste (rangs 78, 110, 134, 167, 181, 418, 433). Extrêmement peu de choses sont connues sur la fonction de la protéine correspondante. Une vérification doit se faire sur l'existence d'un homologue chez l'humain.

Nap1l1 (nucleosome assembly protein 1-like 1): 8 *probesets* sont dans la liste (rangs 130, 173, 227, 607, 623, 648, 942), la protéine correspondante est peu connue et présente 98% homologie avec l'humain. La famille des NAP (*nucleosome assembly protein*) semble être impliquée dans la régulation de l'expression génique.

Gspt1 (G1 to S phase transition 1): la protéine est peu connue mais probablement impliquée dans la régulation du cycle cellulaire et de la croissance cellulaire, et présente 87% homologie avec l'humain. Les 2 *probesets* sont aux rangs 154 et 701.

En fait, beaucoup de gènes suscitent de l'intérêt et on peut encore citer Tmem20, Tmed5, Fibulin1, Rgs3, Thyn1, RAB33A, Bxdc1, Ankrd26. De plus, d'autres gènes qui sont spécifiques des T sans qu'il y ait une surexpression par rapport au contrôle peuvent aussi être d'un grand intérêt comme par exemple Morc (zinc finger, cycle cellulaire, peu connu), Eif3s6 (Int6, site d'intégration du virus MMTV, facteur de transcription) et latrophilin1 (activité de récepteur, transmission de signal, récepteur couplé aux protéines G).

Certains gènes ont été sélectionnés pour être validés par RT-PCR et leurs caractéristiques sont présentées dans la section 4.5.3.3.

4.5.3.1.4. Différences entre T1, T2, T3

Les 3 leucémies T sont hétérogènes et il est nécessaire d'étudier leurs différences. Les *probesets* montrant des variations significatives entre ces 3 leucémies ont été sélectionnés. Les *probesets* ont été classés selon les valeurs décroissantes de l'écart-type et les 3000 premiers *probesets* ont été sélectionnés.

Une analyse SOM ainsi qu'un cluster hiérarchique ont alors été effectués incluant également le contrôle (figure 4.13 A). Ces résultats montrent que T1 (CD4⁺CD8⁺) et T2 (CD4⁺CD8⁺) se ressemblent alors que T3 (CD4⁺CD8⁻) est plus distant. CT (CD4⁺CD8⁺) est plus proche de T1 et T2, mettant T3 encore plus en retrait. La figure 4.13.B illustre une analyse SOM effectuée avec les 3 leucémies T. Cette analyse montre les gènes spécifiques de chaque leucémie par rapport aux autres (T1, T2, T3) et les gènes communs entre T1 et T2 et T2 et T3. Ces résultats n'auraient pas pu être prédits par le phénotype des leucémies.

Le fait que la leucémie T3 soit si différente de T1, T2 et CT nécessite qu'une analyse additionnelle prenant en compte T1 et T2 seulement soit réalisée. Cela va permettre d'approfondir la signature moléculaire de T1 et T2. Le tableau correspondant (appendice K) montre le résultat de cette analyse pour les 50 premiers *probesets*. Une analyse SAM

n'ayant pas pu être effectuée (le minimum d'échantillons dans un groupe étant 3), les *probesets* ont été ordonnés par rapport aux ratios d'expression.

Parmi ces 50 premiers *probesets* présentés, beaucoup sont déjà connus en relation avec la lignée T. Ils attestent de l'état immature des leucémies T1 et T2.

Parmi ceux-là, un gène n'est pas encore relié ni à la lignée T ni aux leucémies T et il serait très intéressant de l'étudier. Il s'agit de *Hdgfrp3* (rangs 16, 25, 46). Il semble être un oncogène potentiel parce que la famille des *Hdgf* (*Hepatoma derived growth factor*) a des caractéristiques mitogéniques. Ce sont des protéines nucléaires capables de lier la chromatine. Il a été choisi pour être validé par RT-PCR.

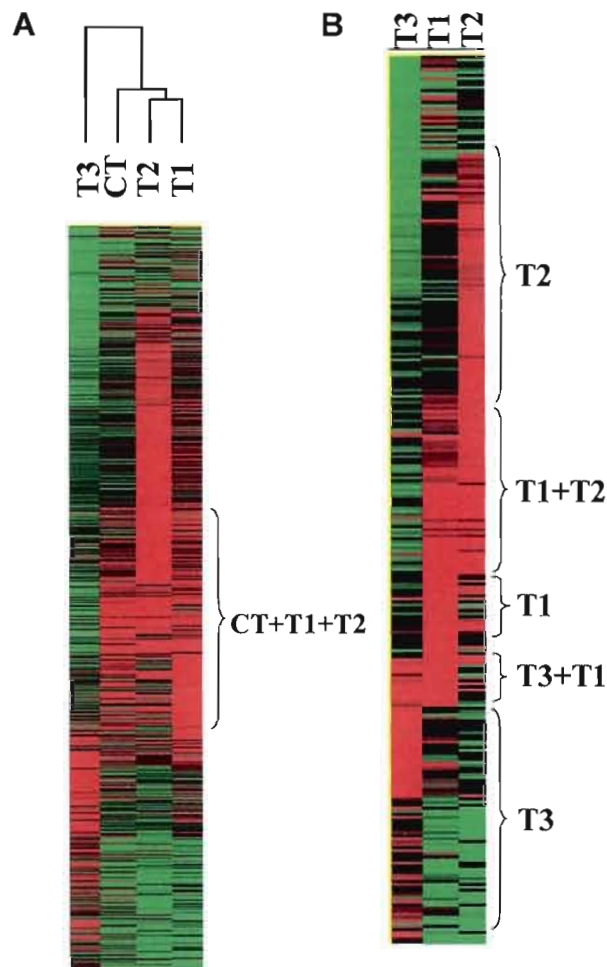


Figure 4.13: Différences entre T1, T2 et T3.(A) *Cluster* hiérarchique réalisé avec T1, T2, T3 et CT. (B) Analyse SOM réalisée avec T1, T2 et T3. La couleur rouge correspond à une surexpression et une couleur verte à une sous-expression par rapport à la moyenne des leucémies.

Beaucoup de résultats intéressants se trouvent aussi au-delà des 50 premiers *probesets* (gènes non connus en relation avec la lignée T). On peut citer par exemple *Elav4* (liaison aux acides nucléiques, régulation de la différenciation neuronale) et *Heyl* (facteur de transcription, liaison à l'ADN) qui pourraient aussi être très vraisemblablement des oncogènes impliqués dans la leucémogénèse de la lignée T (non montré). En conclusion, cette analyse additionnelle apporte de nombreuses données intéressantes qu'il est nécessaire de valider sur un plus grand nombre d'échantillons.

4.5.3.2. Les leucémies B

Les leucémies B induites par le rétrovirus Graffi sont assez hétérogènes similairement aux leucémies T. Leur phénotype commun est CD19⁺B220⁺ mais elles peuvent être B200^{low}, Sca-1⁺ ou⁻. De plus, elles peuvent posséder un réarrangement du BCR ou non et certaines montrent même un réarrangement du TCR (chapitre II). Afin de réaliser les micropuces, 3 leucémies ont été choisies (B1, B2, B3). Elles sont toutes CD19⁺B220⁺ mais B3 est également positive pour Sca-1. Le contrôle (CB) correspond aux cellules B spléniques. Les 3 leucémies ainsi que le contrôle ont été triées avec CD19.

4.5.3.2.1. Vérification des phénotypes

Le tableau 4.19 présenté ci-dessous permet de vérifier les phénotypes.

Tableau 4.19: Vérification des phénotypes des leucémies B

Tableau 4.12: Vérification des phénotypes des leucémies B																
				B220 ⁺ CD19 ⁺ Sca-1	B220 ⁺ CD19 ⁺ Sca-1	B220 ^{low} CD19 ⁺ Sca-1										
	T1	T2	T3	B1	B2	B3	M	E1	E2	E3	Mk1	Mk2	Mk3	T	B	E
Ptpcr (B220)	1.57	1.90	1.29	1.28	0.74	0.52	0.77	-2.89	-4.24	-5.76	0.45	0.98	1.00	1.58	1.29	-0.46
Ptpcr (B220)	1.70	3.44	0.42	1.46	0.67	-0.07	1.92	-2.47	-3.27	-2.76	-1.42	0.10	0.26	1.46	0.51	-1.95
Cd19	-2.77	-3.16	-2.11	5.05	4.82	4.88	-2.24	-1.20	-2.56	-2.48	-2.14	0.98	2.09	-2.48	4.68	-1.36
Ly6a (Sca-1)	-0.87	-3.32	-1.39	-0.65	-0.12	4.42	-0.30	-0.34	-1.82	-2.75	-0.58	1.15	2.05	0.84	3.76	-0.07

Ainsi, l'analyse des micropuces confirme que les 3 leucémies sont CD19⁺, B1 et B2 sont également B220⁺ et B3 B220^{low}. B3 est la seule à exprimer Sca-1. Sca-1 est exprimé par les progéniteurs hématopoïétiques mais, dans le cas de la lignée B, il est en fait un marqueur des cellules B matures (Pflugh et al. 2002). De ce fait, on peut voir que le contrôle B exprime Sca-1 de manière abondante. On peut alors penser que B3, phénotypiquement différente de B1 et B2 est peut-être dans un état plus différencié que les 2 autres leucémies. La plus grande similitude entre B1 et B2 est visible sur le cluster hiérarchique global (figure 4.6.C). Cependant l'analyse en composantes principales (figure 4.5A) montre que les leucémies B sont tout de même plus proches entre elles que les leucémies T.

Il est à remarquer dans le tableau ci-dessus, que Ptpcr (B220), Cd19 et Ly6a (Sca-1) sont aussi exprimés dans d'autres échantillons bien que plus faiblement alors que cela n'avait pas été détecté par l'analyse de cytométrie en flux. Il se peut que l'ARNm soit exprimé mais que les protéines ne soient pas à la surface (restent cytoplasmiques ou pas dans une forme active). Il faut noter pour Ptpcr, qu'il existe des cellules T non conventionnelles extra-thymiques CD3⁺γδCD4⁻CD8⁻B220⁺ dans des souris déficientes ou atteintes de malaria (Bakir et al. 2006). Ces résultats peuvent également être interprétés comme dus à la transformation leucémique qui conduit souvent à l'expression de marqueurs aberrants notamment pour les échantillons Mk.

D'autres antigènes de surface sont exprimés par les échantillons leucémiques B. Il s'agit notamment de Cd22, Cd36, Cd37, Cd38, Cd52, Cd53, Cd79a et b, Cd244 (voir appendice A). Tous ces antigènes exceptés Cd244 (spécifiques des cellules NK) sont connus pour être exprimés par les cellules B (Rutella et al 1999, Erickson et al. 1996, Schwartz-Albiez et al 1988, Oliver et al. 1997, Kubota et al 1990, Olweus et al 1993, Nakajima et Colonna, 2000). Cd36 est un marqueur spécifique des leucémies B car il n'est pas exprimé par les cellules B normales humaines et peu par les cellules B normales murines (Rutella et al. 1999). Son expression est même associée à de faibles chances de guérison (Rutella et al. 1999). Concernant Cd244, il serait intéressant de vérifier son expression dans d'autres échantillons de leucémies murines et humaines car en cas de positivité, il pourrait constituer un nouveau marqueur de leucémies B.

4.5.3.2.2. Gènes surexprimés par les leucémies B

Une analyse SAM a été effectuée sur le même principe que décrit précédemment afin d'identifier les gènes spécifiques des leucémies B. 1636 *probesets* ont passé le critère d'une valeur p inférieure à 0,01 et un FDR inférieur 0,20. Le tableau correspondant (appendice L) présente les 50 premiers *probesets*. À ces valeurs d'expression ont été ajoutées les valeurs du ratio de chaque leucémie B versus le contrôle ('B1-C', 'B2-C', 'B3-C'). Comme le contrôle correspond à des cellules B spléniques matures, les leucémies sont donc très probablement dans un état plus indifférencié que le contrôle. Ainsi les gènes surexprimés par rapport au contrôle comprennent les oncogènes potentiels ainsi que les marqueurs spécifiques des progéniteurs B.

Les 50 premiers *probesets* correspondent à 36 gènes. Parmi ceux-là, de nombreux gènes sont déjà connus et attestent du stade pro-B tardif des cellules leucémiques. Ainsi on retrouve les gènes Vpreb, Igll1 (lambda 5), les gènes codant pour la chaîne lourde des immunoglobulines (Ighg, Igh6) ainsi que Cd79a et b, qui ensemble forment le pre-BCR (LeBien 2000). Les gènes Ebf1, Il7R, Bst1, Foxp1, Enpep, E2A (Tcf2a), E2.2 (Tcf4), témoignent aussi du stade pro-B. Lors du développement des cellules B, la formation du pre-BCR va déclencher plusieurs événements qui vont être: 1) la suppression de l'expression de Rag-1 et Rag-2; 2) la prolifération des pro-B; 3) la réexpression de Rag-1 et Rag-2 permettant l'initiation du réarrangement de la chaîne lourde.

D'autres gènes vont attester de l'état cancéreux des échantillons avec un grand nombre de gènes impliqués dans le cycle cellulaire et qui sont surexprimés par rapport au contrôle: par exemple l'oncogène Myb, les cyclines A2, B1, B2, ainsi que Cdca1, a2, a5, a8 (*cell division cycle associated genes*) (non montré).

Le tableau ci-dessous (tableau 4:20) montre l'exemple de 17 gènes intéressants de par leur patron d'expression et leurs fonctions.

Tableau 4.20: Exemple de gènes d'intérêts pour les leucémies B:

ordre gènes candidats	nom du gène	probesets redondants	rangs des probesets	SAM d-score	p-value	FDR	
1	Pcmt1d1	7	21,66,68,75,83,97,339	19	<0,001	0,01	méthyl-transférase, enzyme de réparation
2	Spred1	6	322,355,778,799,909,921	7,2	<0,001	0,04	inhibition de la voie Ras/Raf/MapK
3	C430010P07Rik	5	132,136,139,539,1358	9,8	<0,001	0,02	rôle inconnu
4	Elovl6	5	30,33,91,105,1079	17	<0,001	0,01	élongation des acides gras
5	Gfra2	4	15,19,296,357	20	<0,001	0,01	récepteur neuronal, transmission du signal
6	Gfra1	4	150,263,1151	9,4	<0,001	0,02	récepteur neuronal, transmission du signal
7	Dyrk2	4	135,192,198,277	9,8	<0,001	0,02	transmission du signal, croissance
8	Ehd2	4	138,144,237,1256	9,7	<0,001	0,02	endocytose, signalisation
9	Rrm2b	4	347,465,555,626	7	<0,001	0,04	réplication, réparation de l'ADN
10	Ndg2	2	16,26	20	<0,001	0,01	signalisation, apoptose
11	Serpinf1	2	18,51	19	<0,001	0,01	anti-angiogénèse
12	Rapgef5	2	35,42	16	<0,001	0,01	développement, signalisation
13	Gpm6a	2	38,41	15	<0,001	0,01	développement neuronal, signalisation
14	Csrp2	1	9	23	<0,001	0,01	développement, différenciation cellulaire
15	Arntl2	1	13	22	<0,001	0,01	facteur de transcription, signalisation
16	Gtse1	1	53	14	<0,001	0,01	cycle cellulaire, apoptose
17	Ckap2	1	1404	4,1	0,007	0,14	cycle cellulaire, apoptose

Ces gènes ne sont pas encore connus en relation avec la lignée B. Ils montrent tous une surexpression par rapport au contrôle. Leur fonction semble en général reliée au processus du cancer (signalisation, cycle cellulaire, apoptose, angiogénèse). Les premiers gènes listés dans le tableau (de 1 à 9) sont représentés par 4 *probesets* et plus, excluant le fait qu'il puisse s'agir de faux-positifs.

Parmi eux, seul Elovl6 (gène candidat 4) ne semble pas avoir une fonction oncogénique. En effet la protéine correspondante semble agir dans l'élongation des chaînes d'acides gras (Moon et al.2001). Cependant cette protéine est très peu connue, et le patron d'expression indique une spécificité certaine pour les leucémies B induites par le rétrovirus Graffi, cela mérite une investigation grâce à d'autres modèles.

Le gène candidat 3, C430010P07Rik, est représenté par 5 *probesets*, et est surexprimé d'un facteur en \log_2 de 3 par rapport au contrôle. Il s'agit d'une opportunité de révéler la fonction de ce gène encore inconnu qui a un rôle certain dans la lignée B que ce soit dans le développement normal ou cancéreux. Gfra1 et 2 (gènes candidats 5 et 6) sont

intéressants car ils codent pour un récepteur membranaire (Jing et al. 1997) qui potentiellement pourrait devenir un nouveau marqueur de surface ainsi que la cible de thérapie avec des anticorps bloquants.

Les autres candidats (Pcmd1, Spred1, Dyrk2, Ehd2, Rrm2b) semblent de par l'annotation fonctionnelle avoir très probablement un rôle oncogénique (DeVry et al.)

Un bon nombre de ces candidats est situé au-delà des 50 premiers *probesets*, montrant que les résultats sont également pertinents au-delà de ce qui est montré dans ce chapitre.

4.5.3.2.3. Gènes communs aux leucémies T et B

Un dernier aspect de l'analyse des leucémies lymphoïdes est la possibilité d'identifier les gènes spécifiques des leucémies T et B versus les autres types de leucémies et les gènes communément surexprimés dans les leucémies T et B par rapport à leur contrôle respectif. 2503 *probesets* ont passé le critère d'une valeur p inférieure à 0,01 et un FDR inférieur à 0,20. Les 100 premiers *probesets* correspondent à 81 gènes. Parmi ceux-là, 40% sont surexprimés par rapport aux contrôles T et B, 49% sont surexprimés dans les leucémies B versus le contrôle B seulement et 11% sont sous-exprimés à la fois dans les leucémies B et T par rapport à leurs contrôles. Le pourcentage (49%) de gènes qui diffèrent entre les leucémies T et B s'explique par le fait que le contrôle T est immature (thymocytes) alors que le contrôle B (cellules B spléniques) est mature. Le tableau correspondant (appendice M) montre les 50 premiers *probesets*.

Le tableau ci-dessous montre l'exemple de 6 gènes intéressants.

Tableau 4.21: Exemples de gènes spécifiques des leucémies T et B

nom des gènes	nombre de probesets	rangs des probesets	SAM d-score	p-value	FDR	T-C*	B-C**	LT***	LB****	
Etv5	3	27, 113, 196	9,9	<0,001	<0,01	+	+	+	+	développement, régulation de la transcription
Edaradd	3	46, 69, 468	8,7	<0,001	<0,01	+	+	+	+	fonction inconnue
Dtl	6	135, 145, 1335, 1584, 2199	6,6	<0,001	<0,01	+	+	+	+	oncogène, développement
Jakmip1	4	4, 26, 71, 239	12,7	<0,001	<0,01	-	+	+	+	signalisation
Ablim1	2	16, 32	10,9	<0,001	<0,01	+	+	+	+	signalisation
Rala	2	35, 966	9,6	<0,001	<0,01	+	+	+	+	signalisation

*: leucémies T- contrôle, **: leucémies B-contrôle, ***: spécifiques des leucémies T, ****: spécifiques des leucémies B

Trois d'entre eux sont surexprimés à la fois dans les leucémies et par rapport aux contrôles (Etv5, Edaradd, Dtl) et 3 sont surexprimés dans les leucémies mais pas par rapport aux contrôles. Etv5 (Erm) est un facteur de transcription Ets et son expression dans les leucémies T n'est pas connue (gène sélectionné pour validation RT-PCR, section 4.5.3.3). Dtl est un oncogène connu notamment dans le cancer du foie (Pan et al. 2006) mais il n'a pas été encore rapporté dans les leucémies. Jakpmip1 est une tyrosine kinase de la famille Jak (gène sélectionné pour validation RT-PCR, section 4.5.3.3). La protéine codée par Ablim1 répond à un signal en agissant sur les protéines du cytosquelette et est impliquée dans la morphogenèse et la différenciation (Roof et al. 1997). Rala est une protéine G similaire à Ras qui voyage entre le cytoplasme et le noyau pour transmettre un signal. Son rôle oncogénique est connu (Lim et al. 2005). Aucun lien avec les lymphocytes n'avait encore été réalisée. MycN et Cdk6 sont les 2 gènes les plus augmentés dans les deux types de leucémies par rapport à leurs contrôles respectifs (moyenne de 5,21 et 3,95) mais ne sont pas spécifiques des leucémies T et B (non montré). Finalement certains marqueurs exprimés par les leucémies T et B sont Cd5, Cd2, Cd26, Cd38 ainsi que le récepteur à l'IL7.

4.5.3.3. Gènes sélectionnés pour validation par RT-PCR

Certains gènes spécifiques des T et des T+B ont été sélectionnés pour être validés par RT-PCR. Il s'agit de 4 gènes spécifiques des leucémies T (Fbln1, Bmp7, Rgs3, Hdgfrp3) et 2 gènes spécifiques des leucémies T et B (Etv5, Jakmip1). Le résumé des caractéristiques de ces gènes est présenté ici.

4.5.3.3.1. Fbln1 (Fibulin 1)

La fibuline 1 est localisée sur le chromosome 15 chez la souris et la protéine correspondante partage 72,5 % d'homologie avec la protéine humaine. La fibuline 1 est une protéine extracellulaire multifonctionnelle. Elle interagit avec les protéines de la matrice extracellulaire étant alors responsable de l'adhésion cellulaire. Elle est donc exprimée par les membranes basales de la plupart des tissus (Roark et al. 1995). Elle est impliquée dans le développement cardiaque, dans la signalisation impliquant des facteurs

de croissance (Perbal et al. 1999). De nombreuses études montrent aussi une relation avec les néoplasies humaines, surtout les cancers du sein et de l'ovaire (Greene et al 2003). Son expression dans ces types de cancer semble être liée à de plus grandes probabilités de guérison et cette protéine semble avoir un effet négatif sur la croissance tumorale (Qing et al. 1997).

Une publication montre que l'expression de Fbln1 par les cellules stromales de la moelle osseuse est régulée par les glucocorticoïdes. De plus, la fibuline 1 se lie à la fibronectine, qui est connue pour réguler la prolifération et la différenciation des cellules souches et progéniteurs hématopoïétiques (Gu et al. 2001). Aucune information ne semble exister quant à l'expression très abondante de Fbln1 dans les leucémies de type T ou les lymphocytes T. Il serait donc intéressant de voir quelles caractéristiques sont apportées aux cellules par son expression abondante. On peut penser à des différences dans la migration, l'adhérence, la prolifération régulée par des signaux externes. Une autre hypothèse peut être une sensibilité différente à la thérapie. En effet, une étude montre que l'agent anti-cancéreux doxorubicine augmente l'expression de fibuline 1 qui à son tour est responsable de la résistance des cellules à cet agent thérapeutique (Pupa et al. 2007).

4.5.3.3.1.2. Bmp7 (bone morphogenetic protein 7)

Les valeurs d'expression en font un gène très spécifique à la lignée T et fortement surexprimé par rapport au contrôle. Deux *probesets* (rangs 83 et 98) sont présents.

Le gène Bmp7 est localisé sur le chromosome 2 chez la souris et la protéine comporte 97% d'homologie avec la protéine humaine. Les protéines BMP (*Bone morphogenetic proteins*) sont des cytokines qui appartiennent à la superfamille TGF β . Les BMPs peuvent contrôler de nombreux processus cellulaires, incluant la prolifération, la différenciation et l'apoptose (Hogan, B.L, 1996). Les différents gènes de la famille BMP, et notamment Bmp7, ont été trouvés exprimés dans différents cancers: cancers de la prostate et du sein, mélanomes (Rothhammer et al. 2005, Doak et al., 2007, Alarmo et al. 2007). Cependant, Bmp7 n'a pas encore été associé avec la leucémie ou la lignée T.

Une étude montre que Bmp7 est régulé par le facteur Lmo4 dans des cellules de cancer du sein (Wang et al. 2007). On peut voir, d'après les résultats des micropuces, que Lmo4 est effectivement plus fortement exprimé dans les leucémies lymphoïdes avec notamment

une très forte expression dans T1 (tableau 4.22). Ainsi, il serait intéressant d'étudier si Lmo4 interagit avec le promoteur de Bmp7 dans les leucémies T induites par le rétrovirus Graffi et dans d'autres modèles plus généraux.

Tableau 4.22: ratios d'expression de Lmo4

	T1	T2	T3	B1	B2	B3	M	E1	E2	E3	Mk1	Mk2	Mk3
Lmo4	2,82	0,74	0,00	1,02	-0,15	0,11	-0,10	-1,61	-1,25	-1,43	-0,56	0,47	-0,08
Lmo4	2,29	0,34	0,97	0,48	0,18	0,53	0,85	-2,04	-1,05	-2,07	-0,14	0,20	-0,54

4.5.3.3.3. Rgs3 (regulator of G-protein signaling 3)

Rgs3 est situé sur le chromosome 4 de la souris et la protéine présente 82% d'homologie avec la protéine humaine. Les protéines RGS jouent un rôle important dans les voies de signalisation des récepteurs couplés aux protéines G. Elles servent de protéines activatrices GTPases pour les protéines G (elles accélèrent l'hydrolyse du GTP). Elles sont donc impliquées dans de nombreuses voies de signalisation. Les RGS sont hautement régulées, faisant d'elles des régulateurs des voies de signalisation dans lesquelles elles sont impliquées. Les RGS semblent être très spécifiques d'un récepteur et d'une protéine G donnés. En général, beaucoup de traitements pharmaceutiques ciblent les récepteurs couplés aux protéines G. Une thérapie ciblée sur les RGS est prometteuse pour cibler de façon plus fine les voies de signalisation impliquée (Riddle et al. 2005). Il existe de nombreux transcrits alternatifs à Rgs3. Il serait intéressant d'élucider le transcrit surexprimé dans les leucémies T induites par le rétrovirus Graffi, bien que beaucoup d'informations sont déjà fournies par les 3 *probesets* d'Affymetrix. Il serait très intéressant d'essayer de déterminer la voie de signalisation ciblée par Rgs3 dans les leucémies T induites par le rétrovirus Graffi et d'élucider en premier lieu quel est le récepteur en amont de la voie de signalisation.

4.5.3.3.4. Hdgfrp3 (hepatoma-derived growth factor, related protein 3)

Ce gène est localisé sur le chromosome 7 chez la souris et la protéine présente 98% d'homologie avec la protéine humaine.

Hdgfrp3 appartient à la famille des Hdgf (*Hepatoma derived growth factor*) qui comprend 6 membres. Hdgf, le premier membre à avoir été découvert, montre des propriétés mitogéniques sur des cellules hépatiques, des fibroblastes, des cellules de muscles lisses et cellules endothéliales (Izumoto et al. 1997). Ce sont des protéines nucléaires capables de lier la chromatine. Hdgfrp3 semble avoir une localisation prédominante dans le cerveau (Abouzied et al. 2004). Sa fonction n'a pas encore été reliée aux leucémies ou à la lignée T. Il est intéressant de voir qu'il semble restreint à 2 phénotypes précis, CD4⁺CD8⁺ et CD4⁺CD8⁺ d'après le résultat des micropuces.

4.5.3.3.5. Etv5 (ets variant gene 5)

Etv5 est situé sur le chromosome 16 chez la souris et présente 95% d'homologie avec l'humain.

Il est un membre de la famille Ets qui sont des facteurs de transcription hélice-boucle-hélice. La famille Ets est particulièrement impliquée dans le système hématopoïétique et par exemple Fli-1 et Pu.1 sont des facteurs Ets. Etv5 est exprimé dans les lymphocytes Th1 CD4⁺ et l'expression est plus élevée lorsque les lymphocytes Th1 sont activés (Cousins et al. 2002). Dans ces cellules, il semble co-opérer avec Stat4 pour réguler l'expression d'interférons (Ouyang et al. 1999). Une étude a aussi trouvé une très forte surexpression de ETV5 dans les leucémies humaines de type B (Korz et al. 2002). Il semble être la cible de la voie de signalisation passant par Ras/Raf-1/MAPK mais les gènes cibles d'Etv5 ne sont pas identifiés (Korz et al. 2002). Il semble avoir les propriétés requises pour être un oncogène dans les leucémies de type B. Ainsi il serait intéressant d'étudier plus en détail les voies de signalisation et les gènes cibles de ce gène dans les leucémies B induites par le rétrovirus Graffi. Les résultats des micropuces suggèrent qu'il pourrait aussi être impliqué dans les leucémies de type T. Il existe plusieurs transcrits alternatifs de Etv5, et il serait intéressant de vérifier si le gène est présent sous des transcrits différents dans les leucémies de type B et T, ce qui semble être suggéré par les résultats des micropuces.

4.5.3.3.6. Jakmip1 (janus kinase and microtubule interacting protein 1)

Jakmip1 est situé sur le chromosome 5 et les protéines, murine et humaine, partagent 60% d'homologie.

Jakmip1 a d'abord été trouvé dans le cerveau où la protéine est capable d'interagir avec le récepteur GABAb et avec des ARNs (Couve et al. 2004). Il a aussi été démontré que Jakmip1 est capable d'interagir avec les Janus kinases et les microtubules. Les Janus kinases sont souvent impliquées dans les voies de signalisation des cellules hématopoïétiques. En fait, Jakmip1 a été trouvé dans un test double-hybride visant à trouver les partenaires d'interaction de la janus kinase Tyk2 dans la lignée cellulaire lymphoïde Jurkat (Steindler et al. 2004). Elle est aussi exprimée dans la lignée cellulaire humaine de lignée B, Daudi (Steindler et al. 2004). L'association de Jakmip1 avec les microtubules fait penser à un rôle dans le réarrangement du cytosquelette, et peut-être un rôle dans la sécrétion de cytokines (Steindler et al. 2004). Il serait donc intéressant de poursuivre l'étude de Jakmip1 en utilisant le modèle Graffi. D'après le résultat des micropuces, la janus kinase Tyk2 n'est pas spécifique des leucémies T mais donne un signal 'présent' dans ces échantillons.

4.6 Résultats partie IV: Validation par RT-PCR

4.6.1 Stratégie générale

Certains gènes ont été sélectionnés afin d'être validés par RT-PCR sur des échantillons *différents* de ceux envoyés pour l'analyse par micropuces. Deux buts ont été associés à cela. Le premier a été de tester le concept. Si, malgré l'hétérogénéité des tumeurs induites par le rétrovirus Graffi, les résultats des micropuces sont valides sur d'autres échantillons, cela signifie que le 'modèle Graffi' est robuste. Si les résultats des micropuces ne sont pas dus au hasard, ils ont donc une plus grande probabilité d'être valides dans d'autres

modèles murins et humains. Le deuxième but est d'établir un lien entre les résultats des micropuces et les protocoles plus traditionnels de biologie: déterminer à partir de quel seuil la PCR en temps réel est nécessaire (la méthode des micropuces étant capable de détecter des variations d'expression très faibles). Cela permet également de connaître l'abondance réelle des transcrits. De plus, il s'agit d'une première étape pour l'étude approfondie de la fonction des gènes sélectionnés.

4.6.2. Matériel et méthodes

Les ARN totaux utilisés ont été obtenus de façon identique aux échantillons envoyés aux micropuces. Le tableau 4.23 montre les phénotypes des cellules leucémiques testés en RT-PCR.

Tableau 4.23: phénotypes des échantillons utilisés en RT-PCR

fig. 4.14 à 4.26		fig. 4.29 à 4.34		phénotypes		triés avec	organe	souris
-	-	1	T#1	T	CD8 ⁺ CD4 ⁻	anti-CD8	thymus	NFS
-	-	2	T#2	T	CD4 ⁺ CD8 ⁺	anti-CD4,CD8	thymus	NFS
-	-	3	T#3	T	CD8 ⁺ CD4 ⁻	anti-CD8	thymus	NFS
1	T#1	4	T#4	T	CD8 ⁺ CD4 ⁻	anti-CD8	rate	NFS
2	T#2	5	T#5	T	CD4 ⁺ CD8 ⁻	anti-CD4	thymus	BALB/c
3	T#3	6	T#6	T	CD4 ⁺ CD8 ⁻	anti-CD4	rate	BALB/c
4	B#1	7	B#1	B	B220 ⁺ Sca-1 ⁺ CD11b ^{low}	anti-B220	ganglions	NFS
5	B#2	8	B#2	B	B220 ⁺ CD19 ⁺	anti-B220	ganglions	NFS
6	B#3	9	B#3	B	B220 ⁺ CD19 ⁺	anti-CD19	ganglions	NFS
7	M	-	-	M	CD11b ⁺	anti-CD11b	rate	NFS
8	E#1	10	E#1	E	Ter119 ⁺ CD71 ⁺	anti-Ter119	rate	NFS
9	E#2	11	E#2	E	Ter119 ⁺ CD71 ⁺	anti-Ter119	rate	NFS
10	E#3	12	E#3	E	Ter119 ^{low} CD71 ⁺	élection négativ	rate	FVB/n
11	Mk#1	13	Mk#1	Mk	CD41 ⁺ cKit ⁺	anti-CD41	rate	NFS
12	Mk#2	14	Mk#2	Mk	CD41 ⁺ cKit ⁺	anti-CD41,cKit	rate	NFS
13	Mk#3	15	Mk#3	Mk	CD41 ⁺ cKit ⁺	anti-CD41	rate	FVB/n
14	Mk#4	-	Mk#4	Mk	CD41 ⁺ cKit ⁺	anti-CD41	rate	NFS

Pour chaque échantillon, la réaction de transcription inverse a été effectuée à partir de 100ng d'ARN total dans un volume de 20µl. (Omniscript, Qiagen, Mississauga, Canada) et à partir d'un oligo-dT. La synthèse du second brin ainsi que la réaction PCR ont été effectuées à partir de 1/10^{ème} de la réaction de RT, dans un volume final de 25µl, selon des conditions standards (Kit PCR, Qiagen, Mississauga, Canada). 1/180^{ème} du volume de RT a été pris pour le contrôle interne β -actine. Le programme PCR appliqué a été le suivant: 95°C 5min 1 fois, 95°C 45s, 57°C (TM-3°C) 45s, 72°C 45s, répétés 25, 27, 29 ou 32 fois selon les conditions, 72°C 10min 1 fois.

Les échantillons ont été migrés sur un gel d'agarose 2% coloré au bromure d'éthidium et le gel a été scanné pour quantification des bandes (Molecular Dynamics Phosphorimager).

Les conditions de RT-PCR avaient été testées et optimisées au préalable à l'aide de 3 gènes (Gata1, 2, et 3) sur plusieurs échantillons leucémiques non triés. Les conditions 'quantitatives' se situaient dans une fourchette comprise entre 24 et 30 cycles en partant de 100ng d'ARN et en prenant ¼ de la réaction RT pour la PCR. Partant de 100ng d'ARN et de 1/10 de la réaction de RT, on peut estimer que la fourchette correspondant aux conditions quantitatives doit être comprise entre 26 et 32 cycles pour des gènes avec un nombre de copies similaires aux gènes testés.

Les régions géniques à amplifier ont été déterminées en observant les régions couvertes par les *probes* Affymetrix correspondant aux gènes amplifiés. En général les *probes* Affymetrix sont situées dans le 5'UTR des gènes mais certains sont également situés dans les exons codants. Une paire d'amorce a été choisie dans le 5'UTR et une autre dans la région codante. La région amplifiée, d'une taille de 250 à 350 paires de base, englobe les régions couvertes par les *probes* Affymetrix. Pour chaque gène testé, la paire d'amorces amplifiant dans le 5'UTR et celle dans les exons codants ont donné le même résultat. Les résultats montrés proviennent le plus souvent de régions amplifiées dans le 5'UTR

4.6.3. Résultats des expériences de RT-PCR

4.6.3.1 Validation de gènes spécifiques des leucémies non-lymphoïdes

4.6.3.1.1. Validation des gènes dans plusieurs échantillons leucémiques

Treize gènes ont été choisis. Le tableau 4.24 en présente la liste.

Tableau 4.24: Gènes non-lymphoïdes sélectionnés

1	Gulp1	Mk
2	Fkbp9	Mk
3	Gucy1a3	Mk
4	Ncf2	B, M, Mk
5	Kit	Mk (B3, M)
6	Gata2	E+Mk
7	Gata1	E+Mk3
8	Snca	E+Mk3
9	Slamf1	E+Mk3
10	Ltpb2	E
11	Cda	E
12	Btbd14a	E
13	Rabgef1	E

Ces gènes ont été présentés dans les sections 4.5.2.4.5 et 4.5.2.5.3. Une intégration rétrovirale a été trouvée à proximité de Gata2 et Kit, ce qui a suscité un intérêt supplémentaire pour valider leurs niveaux d'expression (voir ultérieurement section 4.7).

Les résultats des expériences de RT-PCR pour ces gènes sont présentés de la figure 4.14 à la figure 4.26. Chaque figure est organisée de façon semblable, avec en haut les résultats des micropuces et en bas le résultat des expériences de RT-PCR effectuées sur des échantillons *différents* de ceux des micropuces. Ces échantillons sont numérotés de 1 à 13 ou de 1 à 14 et les phénotypes correspondant sont listés sur le tableau 4.23 (colonnes 1 et 2). Ces échantillons sont nommés T#1 à T#3, B#1 à B#3, etc., pour les différencier des échantillons envoyés pour l'analyse des micropuces, identifiés T1 à T3, B1 à B3, etc..

Pour chaque gène, 29 et 32 cycles de PCR ont été testés, chacune des conditions en duplicata. Le contrôle interne β -actine a été testé à 25 et 29 cycles également en duplicata.

Pour les gènes, les 2 exemplaires du duplicata sont montrés sur les figures pour chaque cycle alors que seulement un exemplaire est montré pour l'actine. L'histogramme présente les résultats de la quantification faite en calculant le ratio des valeurs d'intensité de chaque bande par rapport à l'actine. Chaque barre de l'histogramme représente la moyenne des ratios de 29 et 32 cycles et l'écart-type est calculé entre les 4 expériences. L'écart-type nous donne une estimation des variations expérimentales mais est légèrement biaisé (exagéré) car il y a forcément des variations entre les expériences de 29 et 32 cycles.

Globalement les expériences de RT-PCR confirment très bien les résultats des micropuces. Malgré l'existence de variations expérimentales et l'hétérogénéité des tumeurs, les tendances sont toujours en accord avec les résultats des micropuces. Parmi les échantillons, le numéro 7 (M) donne des résultats constamment faibles pour le contrôle interne β -actine. Cela est probablement dû à une moins bonne qualité de l'échantillon et les résultats correspondant ne peuvent donc pas être interprétés.

Les premières figures (figures 4.14 à 4.16) montrent les gènes spécifiques des leucémies mégacaryoblastiques, *Gulp1*, *Fkbp9* et *Gucy1a3*. Les résultats des RT-PCRs pour ces gènes confirment les résultats des micropuces.

D'après les résultats des RT-PCRs, *Gulp1* montre une très forte expression dans les 3 leucémies mégacaryoblastiques (échantillons 11 à 13) mais également dans E#3 (échantillon 10) et à moindre degré dans 2 leucémies B (échantillons 4 et 6) (figure 4.14).

La quantification montre que *Fkbp9* est exprimé le plus fortement dans Mk#1 (échantillon 11) (figure 4.15). Il est également plus fortement exprimé que la moyenne dans Mk#3, Mk#4 (échantillons 13, 14), et également dans E#3 (échantillon 10).

Figure 4.14 : Gulp1

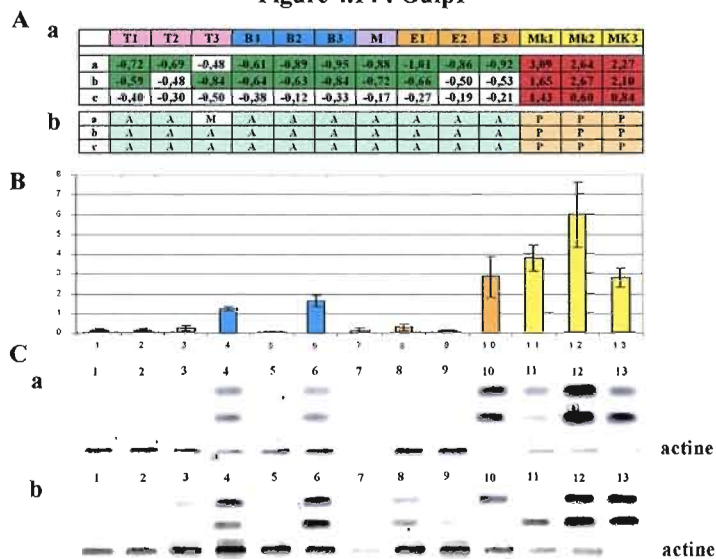


Figure 4.15 : Fkbp9

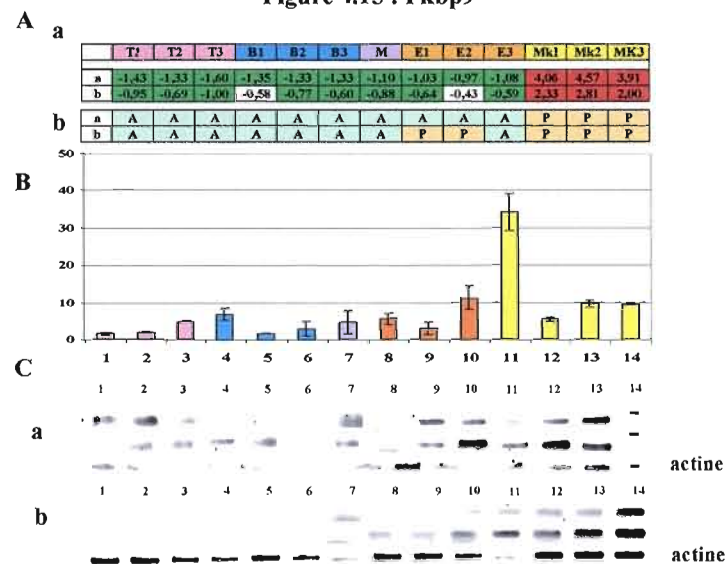


Figure 4.14: Résultats de RT-PCR pour Gulp1. (A) Résultats des micropuces (a) ratios d'expression. Les ratios $> 0,585$ sont colorés en rouge et ceux $< -0,585$ en vert (b) résultats de MAS5.0. (*call*), P: présent (orange), M: médium, A: absent (vert). (B) et (C) RT-PCRs. Echantillons 1-3: leucémies T, 4-6: leucémies B, 7: leucémie M, 8-10: leucémies E, 11-13: leucémies Mk. (B) Histogramme correspondant à la quantification par rapport à β -actine. Code couleur: rose: T, bleu: B, violet: M, orange: E, jaune: Mk. (C) Visualisation des bandes de RT-PCR (a) 29 cycles, (b) 32 cycles.

Figure 4.15: Résultats de RT-PCR pour Fkbp9. (A) Résultats des micropuces (a) ratios d'expression. Les ratios $> 0,585$ sont colorés en rouge et ceux $< -0,585$ en vert (b) résultats de MAS5.0. (call), P: présent (orange), M: médium, A: absent (vert). (B) et (C) RT-PCRs. Echantillons 1-3: leucémies T, 4-6: leucémies B, 7: leucémie M, 8-10: leucémies E, 11-14: leucémies Mk. (B) Histogramme de la quantification par rapport à β -actine. (C) Visualisation des bandes de RT-PCR (a) 29 cycles (b) 32 cycles.

Gucyl3 a un profil très clair. Il est fortement surexprimé dans les 3 leucémies Mk (échantillons 11 à 13) par rapport aux autres échantillons (figure 4.16) et également un peu moins fortement dans E#3 (échantillon 10). Une expression (faible) dans les échantillons 4 (B#1) et 6 (B#3) fait que Gucl3 a globalement un profil très semblable à celui de Gulp1.

Ncf2 est un gène communément surexprimé dans les leucémies de type B, M et Mk d'après les résultats des micropuces (figure 4.17 A). Les résultats de RT-PCR montrent effectivement une forte expression dans les 3 leucémies B (échantillons 4, 5, 6). Ils montrent également une expression supérieure aux autres échantillons dans les 3 leucémies Mk (échantillons 11 à 13), bien qu'à un degré plus faible que pour les leucémies B (figure 4.17 B et C).

Les résultats des RT-PCRs sur Kit concordent également bien avec les résultats des micropuces (figure 4.18). Ce gène est fortement surexprimé dans les 3 leucémies Mk par rapport aux autres échantillons (figure 4.18 B et C). Avec la paire d'amorces utilisées, située dans le 5'UTR du gène, deux fragments de tailles différentes sont amplifiés. Le fragment le plus petit est celui de la taille attendue et c'est celui qui a été quantifié (pointé par des flèches sur la figure 4.18). De façon intéressante, le fragment le plus grand (marqué par des astérisques) apparaît le plus fortement dans les échantillons lymphoïdes mais aussi érythroïdes et n'est pas amplifié dans les leucémies de type mégacaryoblastique.

Le gène Gata2 est spécifique des leucémies érythroïdes et mégacaryoblastiques, avec une surexpression plus forte dans les échantillons Mk d'après les résultats des micropuces (figure 4.19 A). Ceci est confirmé exactement par les expériences de RT-PCR (figure 4.19 B et C).

Figure 4.16 : Gucy1a3

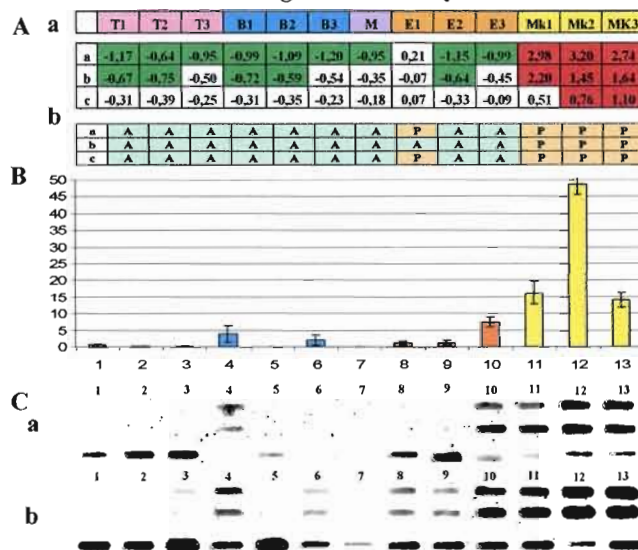


Figure 4.17 : Ncf2

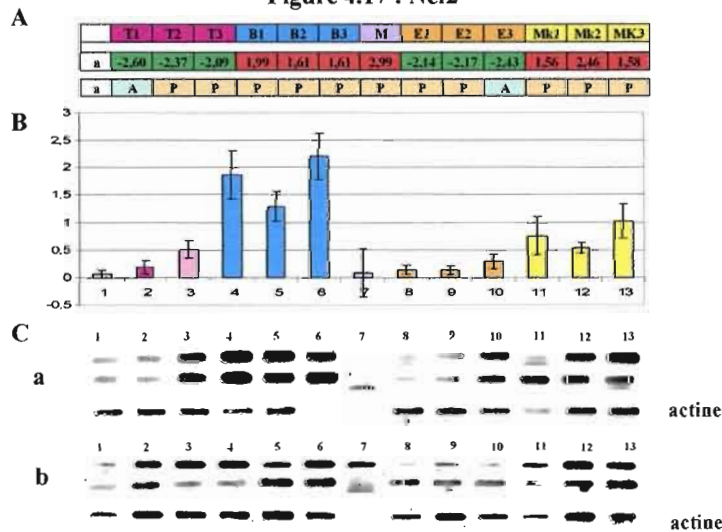


Figure 4.16: Résultats de RT-PCR pour Gucy1a3. (A) Résultats des micropuces (a) ratios d'expression. Les ratios $> 0,585$ sont colorés en rouge et ceux $< -0,585$ en vert (b) résultats de MAS5.0. (call), P: présent (orange), M: médium, A: absent (vert). (B) et (C) RT-PCRs. Echantillons 1-3: leucémies T, 4-6: leucémies B, 7: leucémie M, 8-10: leucémies E, 11-13: leucémies Mk. (B) Histogramme de la quantification par rapport à β -actine. (C) Visualisation des bandes de RT-PCR (a) 29 cycles (b) 32 cycles.

Figure 4.17: Résultats de RT-PCR pour Ncf2. (A) Résultats des micropuces (a) ratios d'expression. Les ratios $> 0,585$ sont colorés en rouge et ceux $< -0,585$ en vert (b) résultats de MAS5.0. (call), P: présent (orange), M: médium, A: absent (vert). (B) et (C): RT-PCRs. Echantillons 1-3: leucémies T, 4-6: leucémies B, 7: leucémie M, 8-10: leucémies E, 11-13: leucémies Mk. (B) Histogramme de la quantification par rapport à β -actine. (C) Visualisation des bandes de RT-PCR (a) 29 cycles (b) 32 cycles.

Le gène Gata1 est le plus spécifique aux leucémies érythroïdes d'après les résultats des micropuces (en accord avec la littérature) et est exprimé à des degrés variables dans les leucémies Mks. Les résultats de RT-PCR montrent une forte spécificité de Gata1 dans les leucémies E et Mks, à des niveaux comparables cependant. Il est à noter que, pour ce gène, aucun fragment n'est amplifié à 29 ou 32 cycles pour les leucémies de type B (échantillons 4 à 6) confirmant les signaux 'absents' et 'medium' de MAS5.0 pour B1, B2 et B3 (= ils ne seraient pas exprimés dans les cellules) (figure 4.20. A, b). Au contraire, T1, T2 et T3, bien que montrant une sous-expression par rapport à la moyenne, ont un signal MAS 5.0 'présent' (= ils sont exprimés dans les cellules) (figure 4.20.A.b). Ainsi des amplifications par RT-PCR sont visibles dans les 3 échantillons, T#1, T#2 et T#3 (de 1 à 3), confirmant précisément les données des micropuces (figure 4.20 B et C).

Les 3 prochains gènes, Snca, Slamf1 et Ltbp2 sont spécifiquement surexprimés dans les leucémies érythroïdes *et également dans Mk3* d'après les résultats des micropuces. Les expériences de RT-PCR confirment leur spécificité pour les leucémies érythroïdes testées (figures 4.21, 4.22, 4.23). Ltbp2 montrant le profil le plus clair. E#3 (échantillon 10) montre la plus forte expression pour Ltbp2, E#1 (échantillon 8) est le plus fort à la fois pour Snca et Slamf1. Slamf1 a une expression plus forte dans les leucémies mégacaryoblastiques par rapport aux 2 autres gènes.

Les 3 derniers gènes testés, Cda, Btbd14a et Rabgef1 ont montré un profil exclusivement érythroïde d'après le résultat des micropuces (figures 4.24, 4.25, 4.26). Les RT-PCR confirment très bien les profils érythroïdes de Cda et Btbd14a.

Cda (figure 4.24) montre le profil le plus clair parmi les 13 gènes testés pour les leucémies non-lymphoïdes. À 29 cycles, aucun fragment n'est amplifié dans les échantillons non érythroïdes, et l'expression est très forte dans les échantillons érythroïdes (échantillons 8, 9, 10). À 32 cycles, on voit une faible expression dans les leucémies mégacaryoblastiques (échantillons 11 à 13).

Figure 4.18 : Kit

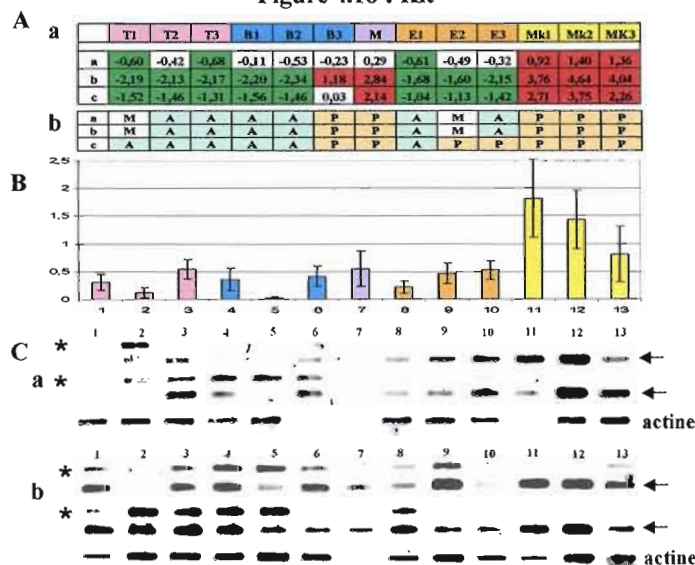


Figure 4.19 : Gata2

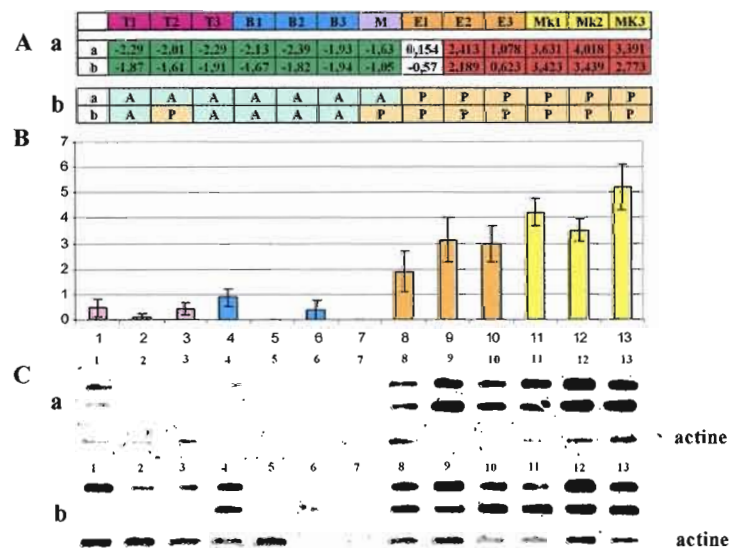


Figure 4.18: Résultats de RT-PCR pour Kit. (A) Résultats des micropuces (a) ratios d'expression. Les ratios $> 0,585$ sont colorés en rouge et ceux $< -0,585$ en vert (b) résultats de MAS5.0. (call), P: présent (orange), M: médium, A: absent (vert). (B) et (C) RT-PCRs. Echantillons 1-3: leucémies T, 4-6: leucémies B, 7: leucémie M, 8-10: leucémies E, 11-13: leucémies Mk. (B) Histogramme correspondant à la quantification par rapport à β -actine.. (C) Visualisation des bandes de RT-PCR (a) 29 cycles (b) 32 cycles.

Figure 4.19: Résultats de RT-PCR pour Gata2. (A) Résultats des micropuces (a) ratios d'expression. Les ratios $> 0,585$ sont colorés en rouge et ceux $< -0,585$ en vert (b) résultats de MAS5.0 (*call*) P: présent (orange), M: médium, A: absent (vert). (B) et (C): RT-PCRs. Echantillons 1-3: leucémies T, 4-6: leucémies B, 7: leucémie M, 8-10: leucémies E, 11-13: leucémies Mk. (B) Histogramme de quantification par rapport à β -actine. (C) Visualisation des bandes de RT-PCR (a) 29 cycles (b) 32 cycles.

Les résultats de Btbd14a confirment également parfaitement les résultats des micropuces (figure 4.25). Premièrement la spécificité érythroïde est confirmée. Deuxièmement, bien que la différence soit faible, il semble que l'expression dans les échantillons mégacaryocytaires est intermédiaire entre les échantillons lymphoïdes et érythroïdes. Cela est également visible dans les résultats des micropuces (figure 4.25.A). On peut voir des ratios fortement négatifs pour les leucémies T, B et M alors que les valeurs sont plus proches de 0 pour les Mk.

Rabgef1, bien que surexprimé spécifiquement dans les leucémies érythroïdes (figure 4.26.A a), a un signal MAS5.0 'présent' dans tous les échantillons (figure 4.26 A b). Cela suggère qu'il est présent de façon relativement abondante dans les autres types cellulaires. Ceci a eu pour conséquence une amplification assez forte dans tous les échantillons testés par RT-PCR à 29 et 32 cycles (figure 4.26 C). La quantification et le calcul des ratios font tout de même ressortir une spécificité érythroïde pour ce gène (figure 4.26 B). On voit également une amplification dans les échantillons mégacaryoblastiques à un degré inférieur. Ceci correspond très bien aux résultats des micropuces et équivaut au cas de Btbd14a. Un test supplémentaire a été effectué à 25 cycles pour Rabgef1 avec l'hypothèse que la spécificité érythroïde peut ainsi ressortir d'une façon plus évidente (figure 4.26 D et E). Cela a donné effectivement de très bons résultats puisque E#1 et E#2 (échantillons 8 et 9) montrent des amplifications fortes alors que des fragments de très faible intensité ont été amplifiés dans les autres échantillons (figure 4.26 E). Les résultats à 25 cycles restent en accord avec ceux obtenus pour les cycles 29 et 32 mais les différences sont plus importantes (figure 4.26 D).

En conclusion, les tests de RT-PCR sur les gènes spécifiques des leucémies non-lymphoïdes confirment bien les résultats des micropuces. On peut remarquer que l'échantillon 10 (E#3) semble porter certaines caractéristiques mégacaryoblastiques (expression de Fkbp9, Gulp1, Gucyl1a3). Quelques échantillons lymphoïdes expriment également assez fortement certains gènes non-lymphoïdes mais toujours à un degré

inférieur aux échantillons non-lymphoïdes. Il existe cependant des variations expérimentales assez grande et un passage en PCR en temps réel améliorerait certainement les résultats.

Figure 4.20 : Gata1

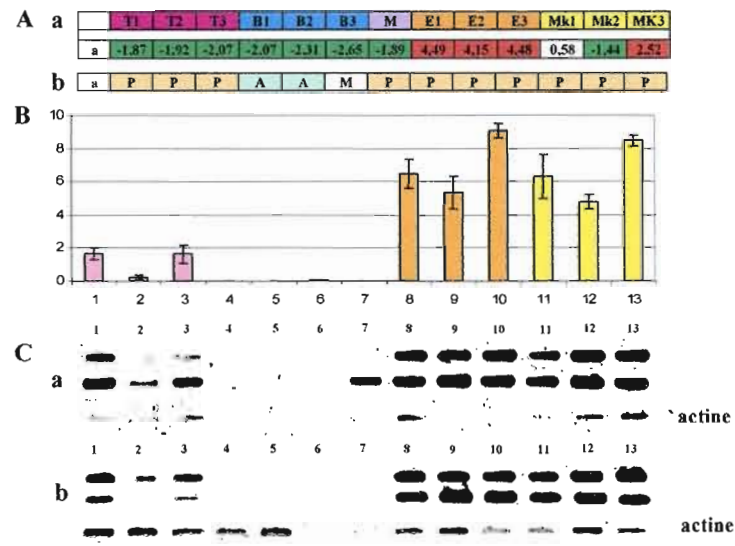


Figure 4.21 : Snca

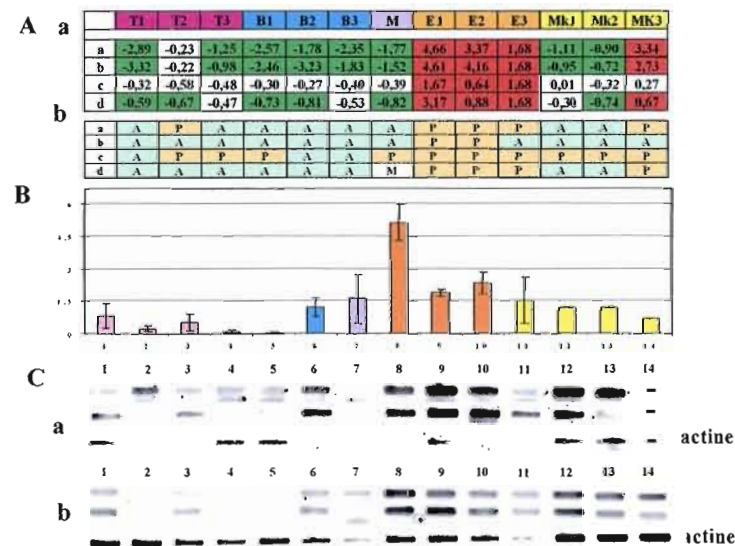


Figure 4.20: Résultats de RT-PCR pour Gata1. (A) Résultats des micropuces (a) ratios d'expression. Les ratios $> 0,585$ sont colorés en rouge et ceux $< -0,585$ en vert (b) résultats de MAS5.0 (*call*), P: présent (orange), M: médium, A: absent (vert). (B) et (C) RT-PCRs. Echantillons 1-3: leucémies T, 4-6: leucémies B, 7: leucémie M, 8-10: leucémies E, 11-13: leucémies Mk. (B) Histogramme de quantification par rapport à β -actine. (C) Visualisation des bandes de RT-PCR (a) 29 cycles (b) 32 cycles.

Figure 4.21: Résultats de RT-PCR pour Snca. (A) Résultats des micropuces (a) ratios d'expression. Les ratios $> 0,585$ sont colorés en rouge et ceux $< -0,585$ en vert (b) résultats de MAS5.0 (*call*), P: présent (orange), M: médium, A: absent (vert). (B) et (C) RT-PCRs. Echantillons 1-3: leucémies T, 4-6: leucémies B, 7: leucémie M, 8-10: leucémies E, 11-14: leucémies Mk. (B) Histogramme de quantification par rapport à β -actine. (C) Visualisation des bandes de RT-PCR (a) 29 cycles (b) 32 cycles.

Les résultats prouvent encore une fois la proximité des leucémies érythroïdes et mégacaryoblastiques par rapport aux leucémies lymphoïdes. Dans beaucoup d'exemples, certains gènes ont une spécificité érythroïde claire et ont une expression mégacaryocytaire intermédiaire entre les échantillons érythroïdes et lymphoïdes (*Snca*, *Ltbp2*, *Slamf1*, *Btbd14a*, *Rabgef1*). Cela est également vrai dans le sens inverse avec une expression des gènes mégacaryocytaires dans E#3 et dans le cas de *Gata2*, une expression très forte pour les Mk et seulement légèrement inférieure pour les échantillons érythroïdes.

4.6.3.1.2. Expression des gènes dans la lignée myéloïde non-leucémique

Les gènes sélectionnés ont également été testés sur des échantillons provenant de cellules de souris saines. Le but était de savoir si ces gènes étaient exprimés dans les cellules non-leucémiques de la lignée myéloïde (non-lymphoïde) et d'obtenir, dans un cas positif, une estimation de l'abondance de ces gènes. Deux essais de RT-PCR ont été effectués (figures 4.27 et 4.28). La figure 4.27 montre les résultats pour les gènes spécifiques des leucémies mégacaryoblastiques alors que la figure 4.28 est ciblée sur les gènes spécifiques des leucémies érythroïdes. Pour ces 2 tests, 3 échantillons non-leucémiques ont été introduits dans les essais. Il s'agit du contrôle érythroïde (CE) constitué d'érythroblastes Ter119⁺, du contrôle myéloïde constitué de progéniteurs myéloïdes CD11b⁺Gr1⁺, ces cellules étant issues de la moelle osseuse de 12 souris saines, ainsi que d'une population *non déterminée* CD41⁺ se trouvant dans la rate de ces souris saines (appelée CMk: contrôle mégacaryocytaire, pour simplifier les explications). Le contrôle érythroïde est l'équivalent de l'échantillon envoyé pour l'analyse par micropuces. La population CD41⁺ issue de la rate n'est malheureusement pas caractérisée par cytométrie en flux. Comme les cellules leucémiques mégacaryoblastiques ont été triées par CD41 à partir de la rate, un contrôle de tri a été fait dans les mêmes conditions à partir de la rate de souris saines. Il a résulté du tri l'obtention d'une population cellulaire bien

définie qui ne colocalise pas tout à fait avec les lymphocytes, de taille pourtant comparable (figure 4.27.H a et b). Elle ne colocalise pas non plus avec les érythrocytes. Il ne peut pas s'agir de plaquettes, les cellules triées étant de taille trop grande et contenant une grande quantité d'ARN. Par comparaison, elles sont plus petites que les cellules mégacaryoblastiques triées à partir des rates leucémiques.

Figure 4.22 : Slamf1

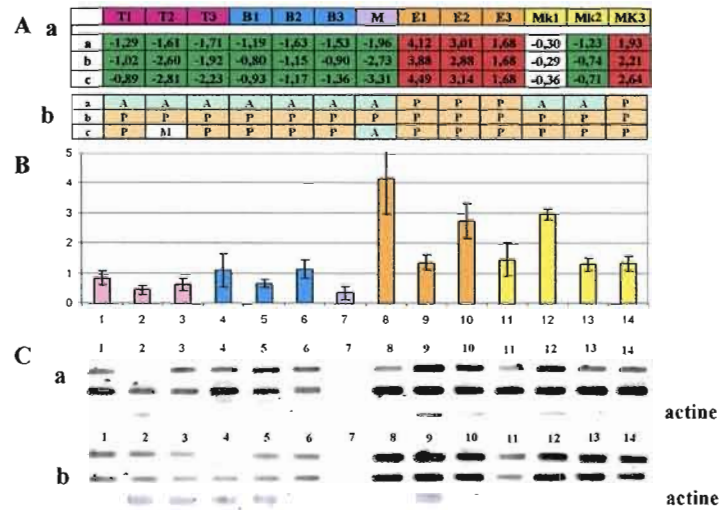


Figure 4.23 : Ltpb2

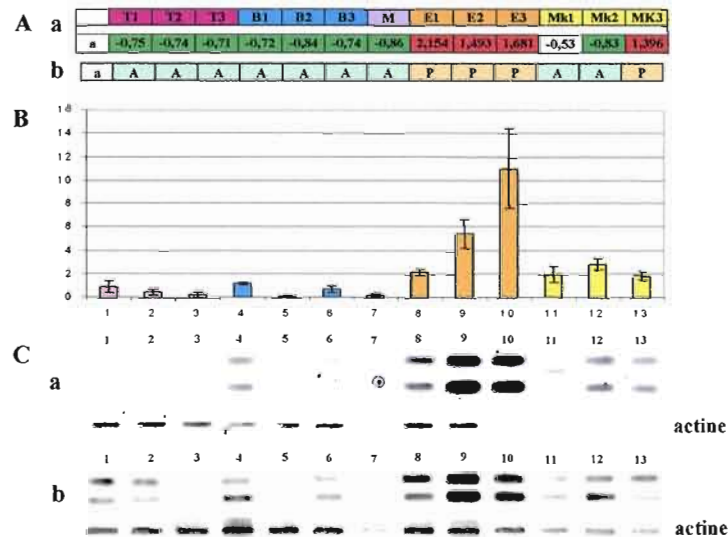


Figure 4.22: Résultats de RT-PCR pour *Slamf1*. (A) Résultats des micropuces (a) ratios d'expression. Les ratios $> 0,585$ sont colorés en rouge et ceux $< -0,585$ en vert, (b) résultats de MAS5.0 (*call*), P: présent (orange), M: médium, A: absent (vert). (B) et (C) RT-PCRs. Echantillons 1-3: leucémies T, 4-6: leucémies B, 7: leucémie M, 8-10: leucémies E, 11-14: leucémies Mk. (B) Histogramme de quantification par rapport à β -actine. (C) Visualisation des bandes de RT-PCR (a) 29 cycles (b) 32 cycles.

Figure 4.23: Résultats de RT-PCR pour *Ltpb2*. (A) Résultats des micropuces (a) ratios d'expression. Les ratios $> 0,585$ sont colorés en rouge et ceux $< -0,585$ en vert (b) résultats de MAS5.0 (*call*), P: présent (orange), M: médium, A: absent (vert). (B) et (C) RT-PCRs. Echantillons 1-3: leucémies T, 4-6: leucémies B, 7: leucémie M, 8-10: leucémies E, 11-13: leucémies Mk. (B) Histogramme de quantification par rapport à β -actine. (C) Visualisation des bandes de RT-PCR (a) 29 cycles, (b) 32 cycles.

Pour simplifier, cette population est appelée échantillon CMk (contrôle mégacaryocytaire) mais rien ne prouve qu'il s'agisse de mégacaryoblastes. Dans une rate normale non triée, 18% des cellules sont positives pour CD41 dont 4% sont CD41⁺Kit⁺ (figure 4.27. H c). Cette population nécessite d'être caractérisée plus en détail.

Pour la figure 4.27, les échantillons testés par RT-PCR ont été: 3 échantillons de leucémies mégacaryoblastiques, 2 échantillons pour la population CD41⁺ (CMk), un échantillon pour le contrôle myéloïde CM, une leucémie érythroïde et un échantillon pour le contrôle érythroïde CE (figure 4.27 A-F). L'attention doit être portée sur la comparaison de l'expression des gènes dans les échantillons des leucémies mégacaryoblastiques (échantillons 1 à 3) par rapport aux échantillons CMk (échantillons 4 et 5), et CM (échantillon 6). On peut observer qu'une amplification est détectée dans les échantillons CMks et CM (échantillons 4, 5, 6) pour tous les gènes testés à l'exception de *Gulp1*. Les résultats de la comparaison sont résumés sur un histogramme (figure 4.27 G). Ainsi, les gènes qui montrent une plus forte expression dans les cellules non leucémiques de la lignée myéloïde (CM, CMK, CE) sont *Kit* et *Ncf2* alors qu'aucune expression n'est détectée pour *Gulp1*. Les autres gènes, *Gucyl3*, *Fkbp9*, et *Gata2*, ont un niveau d'expression intermédiaire. Tous les ratios sont égaux ou inférieurs à 1, suggérant de façon très approximative que ces gènes semblent exprimés moins fortement dans les échantillons non leucémiques par rapport aux 3 leucémies Mk testées (figure 4.27. G).

De façon similaire, des tests de RT-PCR ont été effectués pour les gènes spécifiques des leucémies érythroïdes (figure 4.28). Les échantillons testés ont été: 3 échantillons de leucémies érythroïdes, 2 échantillons pour CE, un échantillon pour une leucémie érythroïde, un échantillon pour la population CD41⁺ (CMk) et un échantillon pour CM

(figure 4.28 A-F). De façon problématique, les échantillons CE ne montrent pas d'amplification visible pour le contrôle interne β -actine même si des amplifications sont observées pour certains gènes testés (C, D, E, F, échantillons 4 et 5). L'échantillon CE contient peu de cellules et est certainement de qualité inférieure. Le fait qu'il y ait une amplification pour les gènes et non pour β -actine n'est pas illogique puisque les conditions de PCR diffèrent: 1/10 de la réaction RT est utilisé pour la PCR pour les gènes alors que 1/180 est utilisé pour β -actine. La normalisation par rapport à β -actine ne peut pas être effectuée. Cependant, les tests de la figure 4.28 ont été réalisés dans une même expérience utilisant les mêmes pools de réactions et les échantillons ont été migrés sur le même gel. Ainsi, il est possible de comparer les valeurs d'expression des échantillons CE entre eux sans calculer le ratio pour connaître *de façon approximative* quels gènes sont plus fortement exprimés par les érythroblastes non leucémiques. Les résultats sont présentés sur l'histogramme de la figure 4.28.G. Snca semble être le plus fortement exprimé dans les échantillons CE, suivi à un degré plus faible par Btbd14a et Rabgef1. Ceci est très bien corrélé avec les résultats des micropuces. Ainsi on voit sur la figure 4.28.H, les ratios d'expression de l'échantillon CE envoyé aux micropuces par rapport aux 3 leucémies érythroïdes E1, E2, E3. Snca est fortement surexprimé dans le contrôle par rapport aux leucémies. Cda et Slamf1 sont les gènes qui montrent l'expression la plus faible dans le contrôle et Btbd14a, Rabgef1 et Ltbp2 ont une expression intermédiaire mais inférieure par rapport aux échantillons leucémiques. Ainsi seul Ltbp2 n'est pas tout à fait corrélé avec le résultat des micropuces. La comparaison par rapport aux échantillons CMk et CM montre que c'est également l'expression de Snca qui est la plus forte, suivie de celles de Rabgef1, Btbd14a, et Slamf1 (figure 4.28).

En conclusion, on peut faire de façon approximative *et non quantitative* un résumé de l'expression des gènes dans chaque type d'échantillon testé. Le tableau 4.25 ci-dessous en présente les résultats.

Tableau 4.25: Résumé de l'expression des gènes non-lymphoïdes

	Mks	Es	CMk	CM	CE	autres
Gucy1a3	+++	-	+--	+--	+--	-
Fkbp9	+++	+-	+-	+-	+-	+-
Gulp1	+++	+-	-	-	-	+-
Gata2	+++	++	+	+-	-	+-
Kit	+++	+-	++	++	+-	+
Ncf2	+++	-	++	++	-	+
Ltbp2	+-	+++	+-	+-	-	+-
Cda	-	+++	-	-	-	-
Snca	+	+++	++	++	++	+
Btbd14a	+	+++	+	+	+	-
Rabgef1	+	+++	+	+	+	+
Slamf1	+	+++	-	-	-	+

ordre décroissant de la force de l'expression: +++, ++, +, +-, +--, -

autres= échantillons leucémiques lymphoïdes

Certains gènes, bien que spécifiques des leucémies mégacaryoblastiques ou érythroïdes selon les cas, ont aussi une expression assez forte dans les échantillons non-leucémiques ou dans les échantillons leucémiques non-lymphoïdes. Ces gènes plus ubiquitaires sont dans l'ordre décroissant: Snca, Rabgef1 et Kit. Inversement, certains sont plus spécifiques et de l'état leucémique et des leucémies mégacaryoblastiques ou érythroïdes. Cda est de loin le plus spécifique avec une expression restreinte aux échantillons leucémiques érythroïdes, suivi par Slamf1, Gulp1 et Gucy1a3.

Ces analyses de RT-PCR permettent de caractériser un peu mieux la population splénique CD41⁺. L'expression des gènes semble corrélée à celle des échantillons CE et CM. De façon très intéressante, elle exprime très fortement Kit (alors que CE ne l'exprime pas). Cette population, si elle est CD41⁺Kit⁺, serait-elle un progéniteur mégacaryoblastique? Les données peu nombreuses semblent indiquer que la mégacaryopoïèse s'effectue principalement dans la moelle osseuse chez la souris (Long and Williams 1982). Cependant, il est bien connu que la rate est, chez la souris, un organe

Figure 4.24 : Cda

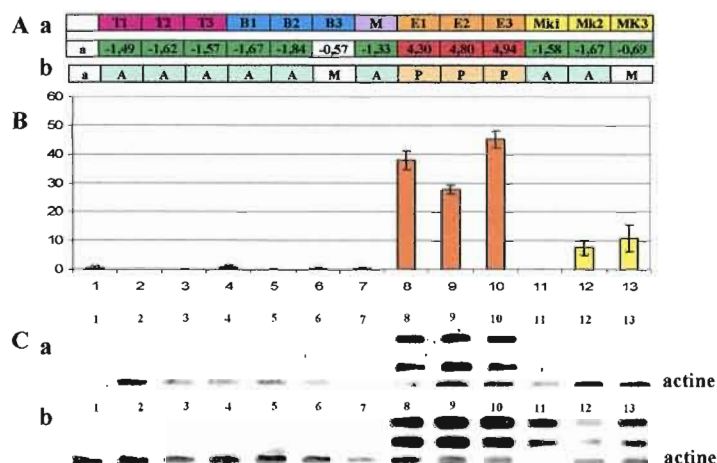


Figure 4.25 : Btbd14a

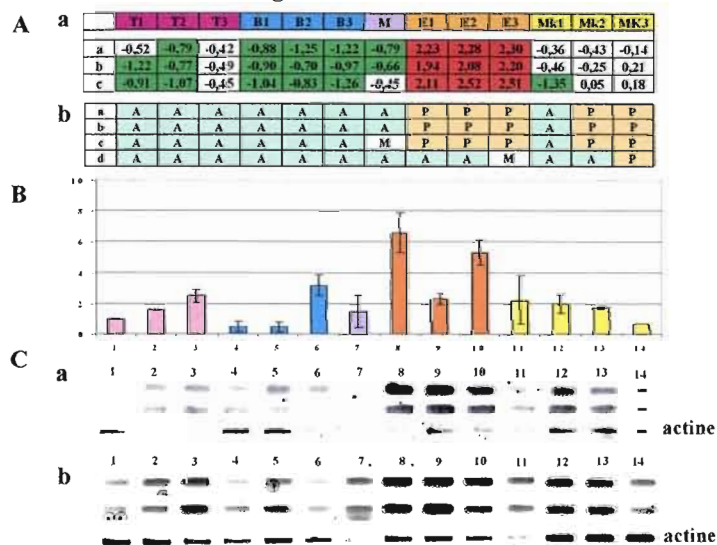


Figure 4.24: Résultats de RT-PCR pour Cda. (A) Résultats des micropuces (a) ratios d'expression. Les ratios $> 0,585$ sont colorés en rouge et ceux $< -0,585$ en vert (b) résultats de MAS5.0 (*call*), P: présent (orange), M: médium, A: absent (vert). (B) et (C) RT-PCRs. Echantillons 1-3: leucémies T, 4-6: leucémies B, 7: leucémie M, 8-10: leucémies E, 11-13: leucémies Mk. (B) Histogramme de quantification par rapport à β -actine. (C) Visualisation des bandes de RT-PCR (a) 29 cycles (b) 32 cycles.

Figure 4.25: Résultats de RT-PCR pour Btbd14a. (A) Résultats des micropuces (a) ratios d'expression. Les ratios $> 0,585$ sont colorés en rouge et ceux $< -0,585$ en vert (b) résultats de MAS5.0 (*call*), P: présent (orange), M: médium, A: absent (vert). (B) et (C): RT-PCRs. Echantillons 1-3: leucémies T, 4-6: leucémies B, 7: leucémie M, 8-10: leucémies E, 11-14: leucémies Mk. (B) Histogramme de quantification par rapport à β -actine. (C) Visualisation des bandes de RT-PCR (a) 29 cycles (b) 32 cycles.

hématopoïétique et est notamment le siège d'une érythropoïèse extra-médullaire. Comme les 2 lignées sont très proches, il ne serait pas illogique qu'un site secondaire de la mégacaryopoïèse se situe également dans la rate. Par exemple, une étude prouve que la rate est le site principal de la mégacaryopoïèse après la transplantation de cellules souches hématopoïétiques (Slayton et al. 2002). Cependant aucune preuve n'est apportée ici. La population splénique CD41⁺ reste indéterminée et une caractérisation approfondie en cytométrie de flux est nécessaire.

Figure 4.26 : Rabgef1

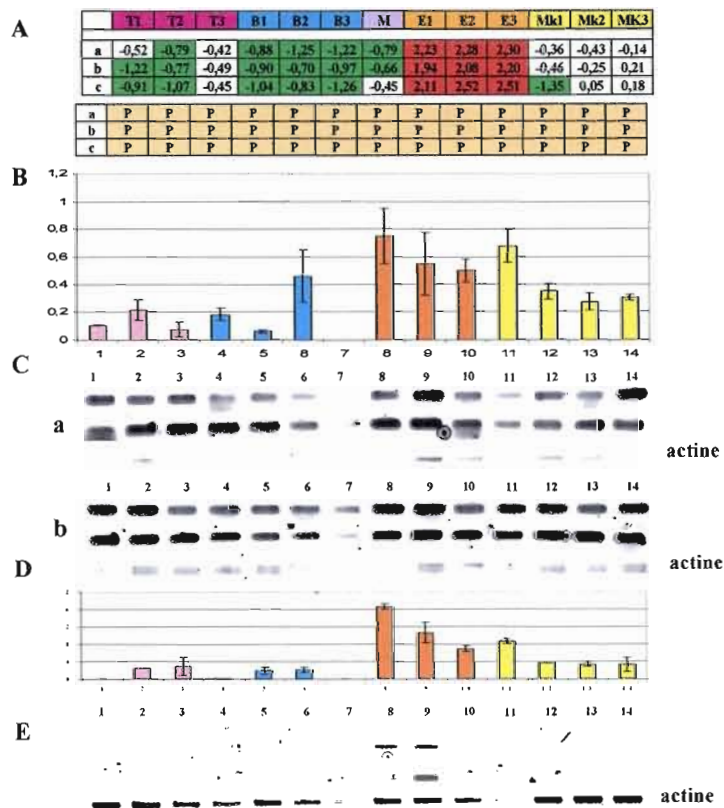


Figure 4.26: Résultats de RT-PCR pour Rabgef1. (A) Résultats des micropuces. (a) ratios d'expression. Les ratios $> 0,585$ sont colorés en rouge et ceux $< -0,585$ en vert (b) résultats de MAS5.0 (*call*), P: présent (orange), M: médium, A: absent (vert). (B) (C) (D) (E) RT-PCRs. Echantillons 1-3: leucémies T, 4-6: leucémies B, 7: leucémie M, 8-10: leucémies E, 11-14: leucémies Mk. (B) Histogramme de quantification par rapport à β -actine. (C) Visualisation des bandes de RT-PCR (a) 29 cycles (b) 32 cycles. (D) Histogramme correspondant à la quantification de la RT-PCR à 25 cycles. (E) Visualisation des bandes de RT-PCR à 25 cycles.

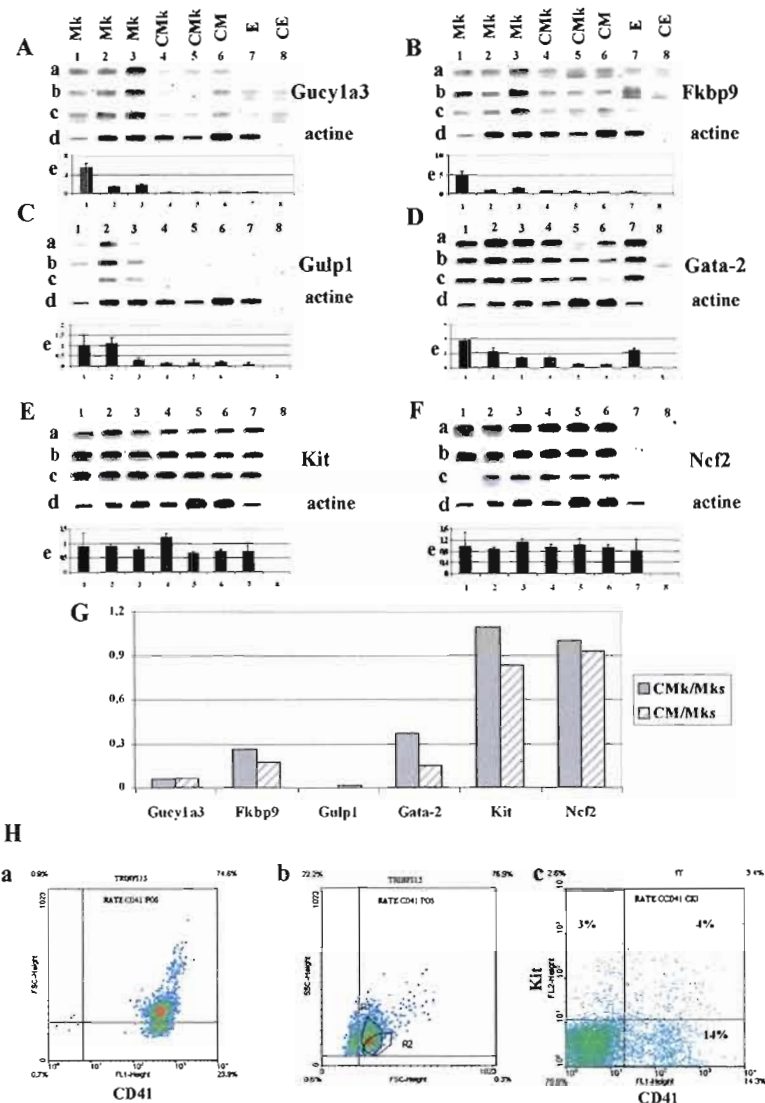


Figure 4.27: Expression des gènes spécifiques des leucémies mégacaryoblastiques dans des échantillons non leucémiques.

Échantillons 1 à 3: leucémies mégacaryoblastiques Mk#1, Mk#2, Mk#3; échantillons 4, 5: CMks, cellules spléniques non leucémiques CD41⁺; échantillon 6: CM, contrôle myéloïde CD11b⁺Gr1⁺; échantillon 7: leucémie érythroïde E#3; échantillon 8: CE, contrôle érythroïde Ter119⁺.

(A) à (F) Résultat des RT-PCRs. (A) Gucy1a3 (B) Fkbp9 (C) Gulp1 (D) Gata2 (E) Kit (F) Ncf2. (a)-(c) triplicata pour le gène testé (d) β -actine (e) histogramme de quantification par rapport à β -actine. (G) Histogramme de quantification des échantillons CMk (barres grises) et CM (barres hachurées) versus la moyenne des échantillons Mk. (H) Analyse en cytométrie de flux des échantillons CMk après le tri (a) Scatter plot sur FSH, F11, indiquant que 100% des cellules sont fluorescentes en F11 (CD41⁺) (b) Scatter plot sur FSH, SSH montrant la taille des cellules CD41⁺. La région R2 montre l'emplacement des lymphocytes (c) analyse d'une rate normale non triée avec les marqueurs spécifiques de CD41 et Kit.

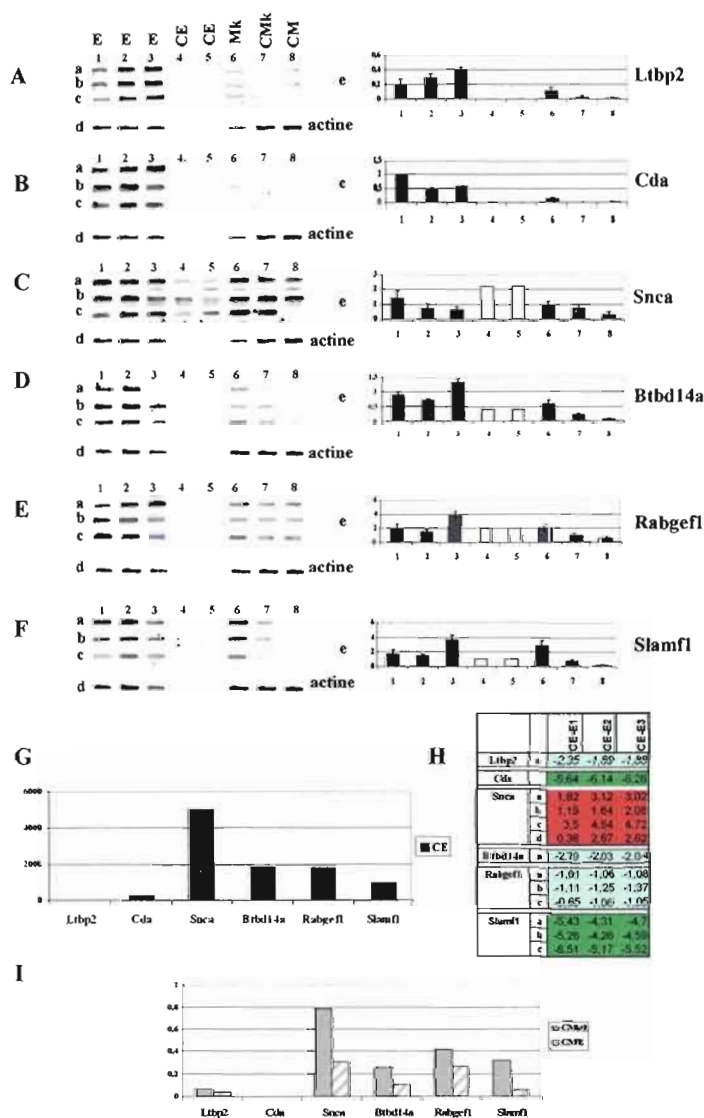


Figure 4.28: Expression des gènes spécifiques des leucémies érythroïdes dans des échantillons non leucémiques.

Échantillons 1 à 3: leucémies érythroïdes E#1, E#2, E#3 ; échantillons 4,5: CE = contrôle érythroïde Ter119⁺ ; échantillon 6: leucémie mégacaryocytaire Mk#1 ; échantillon 7: CMk, cellules spléniques non leucémiques CD41⁺ ; échantillon 8: CM, contrôle myéloïde CD11b⁺Gr1⁺.

(A) à (F) Résultats des RT-PCRs (A) Ltbp2 (B) Cda ; (C) Snca ; (D) Btbd14a ; (E) Rabgef1 ; (F) Slamf1. (a)-(c) triplicata pour le gène testé (d) β -actine (e) histogramme de quantification par rapport à β -actine. (G) Histogramme représentant les valeurs des échantillons CE (4 et 5) pour chaque gène testé (H) Résultats des micropuces montrant le

4.6.3.2. Validation de gènes spécifiques des leucémies lymphoïdes

4.6.3.2.1. Gènes spécifiques des T et des B

Six gènes ont été testés et la liste est présentée sur le tableau 4.26. Quatre gènes sont spécifiques des leucémies T uniquement et 2 sont spécifiques à la fois des leucémies T et B.

Tableau 4.26: Gènes lymphoïdes sélectionnés

1	Fbln1	T
2	Rgs3	T
3	Bmp7	T
4	Hdgfrp3	T1+T2
5	Etv5	T+B
6	Jakmip1	T+B

Les résultats des expériences de RT-PCR pour ces gènes sont présentés de la figure 4.29 à la figure 4.34. La description des échantillons utilisés est présentée sur le tableau 4.23. Les RT-PCRs ont été effectuées une première fois à 25 cycles avec les échantillons de 4 à 15 (identiques à ceux utilisés pour les gènes spécifiques des leucémies non-lymphoïdes, excepté l'omission de l'échantillon M). Le nombre de cycles a été choisi par rapport aux résultats des RT-PCR précédentes, notamment avec le résultat pour Rabgef1 qui était meilleur à 25 cycles (figure 4.26). Les résultats à 25 cycles étaient satisfaisants mais l'intensité globale était faible, un autre essai a alors été effectué à 28 cycles. De plus, 3 échantillons supplémentaires de leucémies T ont été rajoutés. Ainsi les 3 phénotypes, $CD4^+CD8^+$, $CD4^+CD8^-$ et $CD4^-CD8^+$ sont représentés dans les tests de RT-PCR. Les phénotypes des échantillons qui ont été envoyés pour l'analyse par micropuces étaient $CD4^+CD8^+$ pour T1, $CD4^-CD8^+$ pour T2 et $CD4^+CD8^-$ pour T3. Ainsi, les échantillons testés en RT-PCR T#1, T#3 et T#4 ont le même phénotype que T2, T#2 est l'équivalent de T1 et T#5 et T#6 sont l'équivalent de T3. Il est à noter que T#1 et T#2 proviennent d'une même leucémie qui contenait 2 populations, l'une $CD4^+CD8^+$ et l'autre $CD8^+CD4^-$.

De façon similaire à la section précédente, les 6 figures sont organisées sur un même modèle. Les résultats des micropuces sont d'abord présentés (parties A des figures), suivi de l'histogramme de quantification des RT-PCRs (parties B des figures) et des résultats des gels d'agarose des RT-PCRs (parties C des figures). Les parties D des figures

représentent les résultats des micropuces des leucémies T et/ou B par rapport à leur échantillon contrôle respectif.

Figure 4.29 Fbln1 (T)

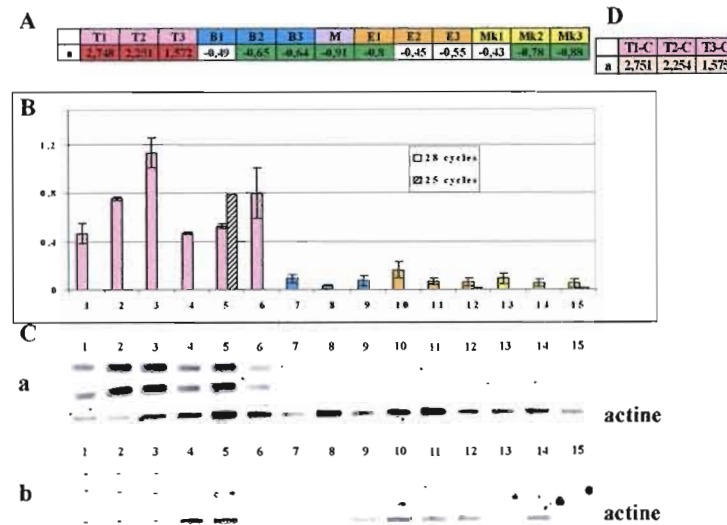


Figure 4.30 Bmp7 (T)

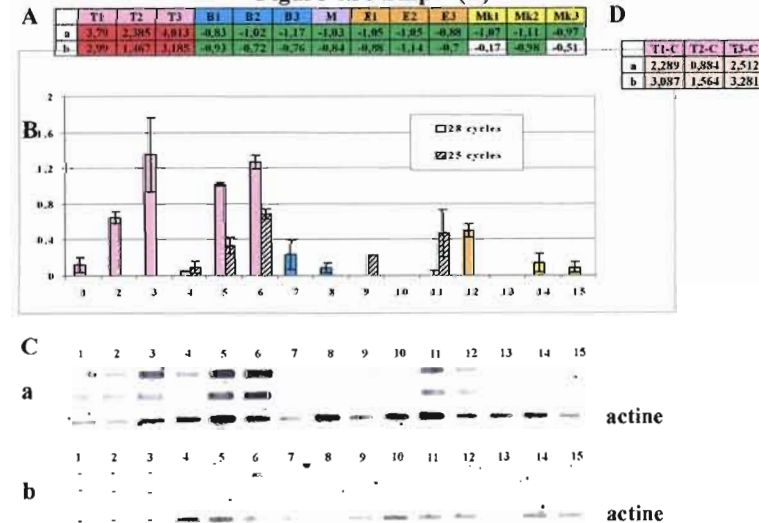


Figure 4.29: Résultats de la RT-PCR pour Fbln1. (A) Résultats des micropuces. Les ratios supérieurs à 0,585 sont colorés en rouge et ceux inférieurs à -0,585 en vert. (B) et (C) RT-PCR. Echantillons 1-6: leucémies T, 7-9: leucémies B, 10-12: leucémies E, 13-15: leucémies Mk. (B) Histogramme de quantification par rapport à β -actine. Les barres en couleur présentent les résultats à 28 cycles, les barres hachurées présentent les résultats à 25 cycles. (C) Visualisation des bandes de RT-PCR (a) 28 cycles; (b) 25 cycles. (D) résultats des micropuces montrant le ratio du contrôle T par rapport à chaque leucémie T. En vert clair: ratio inférieur à 0,585; en orange: ratio supérieur à 0,585.

Figure 4.30: Résultats de la RT-PCR pour Bmp7. (A) Résultats des micropuces. Les ratios supérieurs à 0,585 sont colorés en rouge et ceux inférieurs à -0,585 en vert. (B) et (C) RT-PCR. Echantillons 1-6: leucémies T, 7-9: leucémies B, 10-12: leucémies E, 13-15: leucémies Mk. (B) Histogramme de quantification par rapport à β -actine. Les barres en couleur présentent les résultats à 28 cycles, les barres hachurées présentent les résultats à 25 cycles. (C) Visualisation des bandes de RT-PCR (a) 28 cycles; (b) 25 cycles. (D) résultats des micropuces montrant le ratio du contrôle T par rapport à chaque leucémie T. En vert clair: ratio inférieur à 0,585; en orange: ratio supérieur à 0,585.

Pour les 6 gènes testés, les résultats des RT-PCRs confirment précisément les résultats des micropuces.

Les 3 premiers gènes testés, Fbln1, Rgs3, Bmp7 (figures 4.29 à 4.31) sont spécifiques des 3 leucémies de type T d'après le résultat des micropuces et surexprimés par rapport au contrôle.

Les résultats de RT-PCR montrent que Fbln1 est fortement exprimé dans les 6 leucémies T testées et aucune amplification n'est détectée dans les autres échantillons (figure 4.29 B et C).

Les résultats de RT-PCR et des micropuces sont également très bien corrélés pour Bmp7, avec une expression restreinte aux échantillons de type T (figure 4.30). Parmi les échantillons, on note cependant des variations, avec 2 échantillons $CD4^+CD8^+$ sur 3 exprimant très faiblement le gène (échantillons T#1 et T#4). Une amplification est détectée dans 2 leucémies de type érythroïde (échantillons 11 et 12).

Le gène Rgs3 confirme également son profil spécifique des leucémies T (figure 4.31). Son profil ressort mieux à 25 cycles (figure 4.31.b). Il existe des variations d'expression entre les échantillons T. Il apparaît plus ubiquitaire avec des amplifications, même faibles, détectées dans chaque échantillon à 28 cycles. Cette ubiquité était apparente sur les résultats de MAS 5.0 (non montrés). Un échantillon mégacaryocytaire (échantillon 15) montre une expression de Rgs3 plus élevée.

Hdgfrp3 apparaît spécifique des leucémies T1 et T2 ($CD4^+CD8^+$ et $CD4^-CD8^+$) d'après le résultat des micropuces (figure 4.32 A). Cette spécificité est confirmée avec les tests des RT-PCRs. En effet une forte amplification est détectée dans les échantillons T#1 à T#4 qui sont $CD4^+CD8^+$ et $CD4^-CD8^+$ et aucune amplification n'est détectée dans les échantillons T#5 et T#6 qui sont $CD4^+CD8^-$ comme T (figure 4.32 B et C).

Figure 4.31 Rgs3 (T)

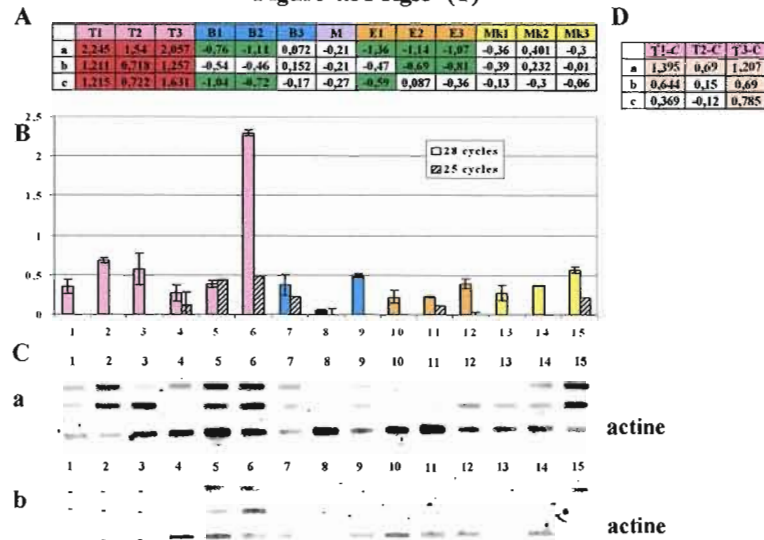


Figure 4.32 Hdgfrp3 (T1+T2)

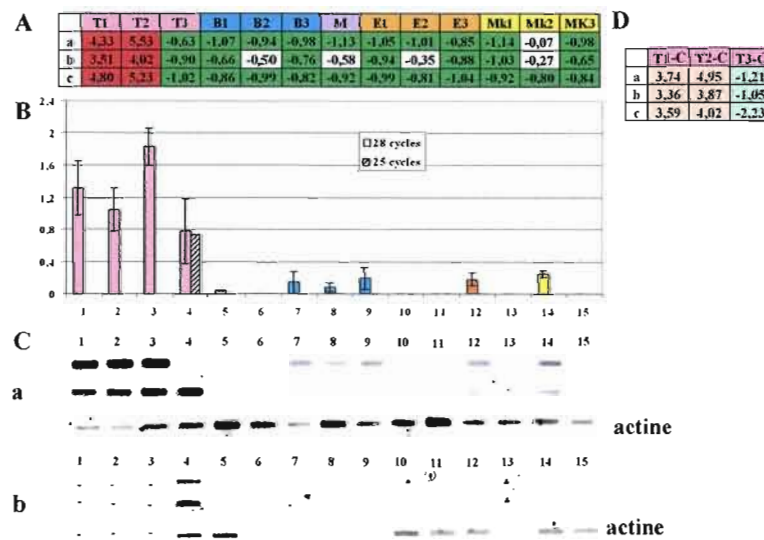


Figure 4.31: Résultats de la RT-PCR pour Rgs3. (A) Résultats des micropuces. Les ratios supérieurs à 0,585 sont colorés en rouge et ceux inférieurs à -0,585 en vert. (B) et (C) RT-PCR. Echantillons 1-6: leucémies T, 7-9: leucémies B, 10-12: leucémies E, 13-15: leucémies Mk. (B): Histogramme de quantification par rapport à β -actine. Les barres en couleur présentent les résultats à 28 cycles, les barres hachurées présentent les résultats à 25 cycles. (C) Visualisation des bandes de RT-PCR (a) 28 cycles; (b) 25 cycles. (D) résultats des micropuces montrant le ratio du contrôle T par rapport à chaque leucémie T. En vert clair: ratio inférieur à 0,585; en orange: ratio supérieur à 0,585.

Figure 4.32: Résultats de la RT-PCR pour *Hdgfrp3*. (A) Résultats des micropuces. Les ratios supérieurs à 0,585 sont colorés en rouge et ceux inférieurs à -0,585 en vert. (B) et (C) RT-PCR. Echantillons 1-6: leucémies T, 7-9: leucémies B, 10-12: leucémies E, 13-15: leucémies Mk. (B) Histogramme de quantification par rapport à β -actine. Les barres en couleur présentent les résultats à 28 cycles, les barres hachurées présentent les résultats à 25 cycles. (C) Visualisation des bandes de RT-PCR (a) 28 cycles (b) 25 cycles. (D) résultats des micropuces montrant le ratio du contrôle T par rapport à chaque leucémie T. En vert clair: ratio inférieur à 0,585 ; en orange: ratio supérieur à 0,585.

Les 2 gènes suivants, *Etv5* et *Jakmip1*, sont spécifiques à la fois des leucémies de type T et B d'après les résultats des micropuces (figures 4.33 et 4.34). *Etv5* est un peu plus fortement exprimé dans les leucémies B alors que *Jakmip1* l'est de façon équivalente dans les T et B, ceci toujours à partir des résultats des micropuces. Les RT-PCRs confirment très bien ce profil avec des amplifications marquées dans les échantillons lymphoïdes pour *Etv5* et *Jakmip1*.

4.6.3.2.2. Validation de gènes spécifiques des leucémies B

Cette partie a été effectuée par Cyndia Charfi au cours de sa maîtrise et est continuée au cours de son doctorat. La figure 4.35 montre l'exemple de 2 gènes spécifiques des leucémies B: *Arntl2* (appendice L, rang 13) et *Gpm6a* (appendice L, rang 38). Les résultats des RT-PCR montrent qu'ils sont bien surexprimés par rapport aux leucémies T et par rapport aux contrôles T et B.

D'autres gènes spécifiques des leucémies B ont été testés et confirmés par Cyndia Charfi (travail de doctorat en cours). Il s'agit de: *Bfsp2* (appendice L, rang 3), *Fmn2* (appendice L, rang 1), *Gfra2* (appendice L, rang 15), et *Gpm6b*.

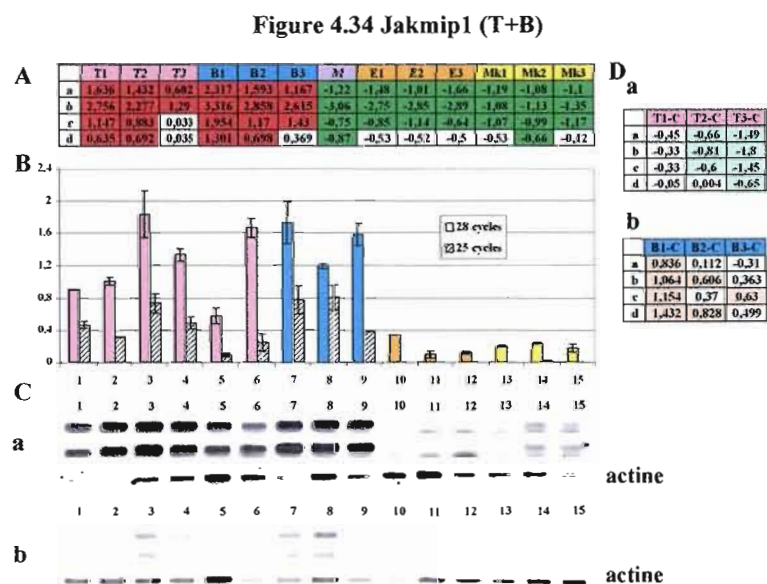
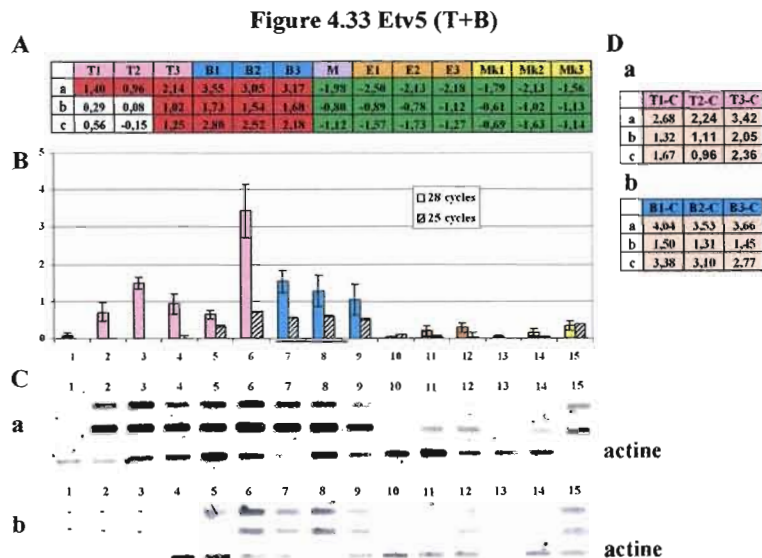


Figure 4.33: Résultats de la RT-PCR pour Etv5. (A) Résultats des micropuces. Les ratios supérieurs à 0,585 sont colorés en rouge et ceux inférieurs à -0,585 en vert. (B) et (C) RT-PCR. Echantillons 1-6: leucémies T, 7-9: leucémies B, 10-12: leucémies E, 13-15: leucémies Mk. B: Histogramme de quantification par rapport à β -actine. Les barres en couleur présentent les résultats à 28 cycles, les barres hachurées présentent les résultats à 25 cycles. (C) Visualisation des bandes de RT-PCR (a) 28 cycles (b) 25 cycles. (D) résultats des micropuces montrant le ratio des contrôles par rapport à chaque leucémie lymphoïde. En vert clair: ratio inférieur à 0,585; en orange: ratio supérieur à 0,585. (a) leucémies et contrôle T (b) leucémies et contrôle B.

Figure 4.34: Résultats de la RT-PCR pour Jakmip1. (A). Résultats des micropuces. Les ratios supérieurs à 0,585 sont colorés en rouge et ceux inférieurs à -0,585 en vert. (B) et (C) RT-PCR. Echantillons 1-6: leucémies T, 7-9: leucémies B, 10-12: leucémies E, 13-15: leucémies Mk. (B): Histogramme de quantification par rapport à β -actine. Les barres en couleur présentent les résultats à 28 cycles, les barres hachurées présentent les résultats à 25 cycles. (C) Visualisation des bandes de RT-PCR (a) 28 cycles; (b) 25 cycles. (D) résultats des micropuces montrant le ratio des contrôles par rapport à chaque leucémie lymphoïde. En vert clair: ratio inférieur à 0,585; en orange : ratio supérieur à 0,585. (a) leucémies et contrôle T; (b) leucémies et contrôle B.

En conclusion, les gènes sélectionnés correspondant à chaque type de leucémie analysée ont été globalement confirmés par RT-PCR. Cela valide l'expérience dans le modèle Graffi et montre que l'hétérogénéité des tumeurs bien que réelle n'a pas freiné la possibilité d'obtenir les gènes les plus spécifiques de chaque type de leucémie. La diversité des leucémies induites par le rétrovirus Graffi peut être utilisée comme un avantage du modèle. Les perspectives seront notamment de confirmer la spécificité de ces gènes dans d'autres modèles notamment chez l'humain (voir Discussion section 4.8). Les gènes sélectionnés présentent des variations d'expression très grandes entre les types de leucémies. Pour d'autres gènes d'intérêt présentant des variations plus faibles, un passage en PCR en temps réel est nécessaire. Les expériences de RT-PCR ont permis de faire la relation entre les valeurs d'expression des micropuces, provenant de calculs biostatistiques compliqués et l'abondance relative des transcrits dans les tissus testés de façon plus classique et plus palpable par RT-PCR.

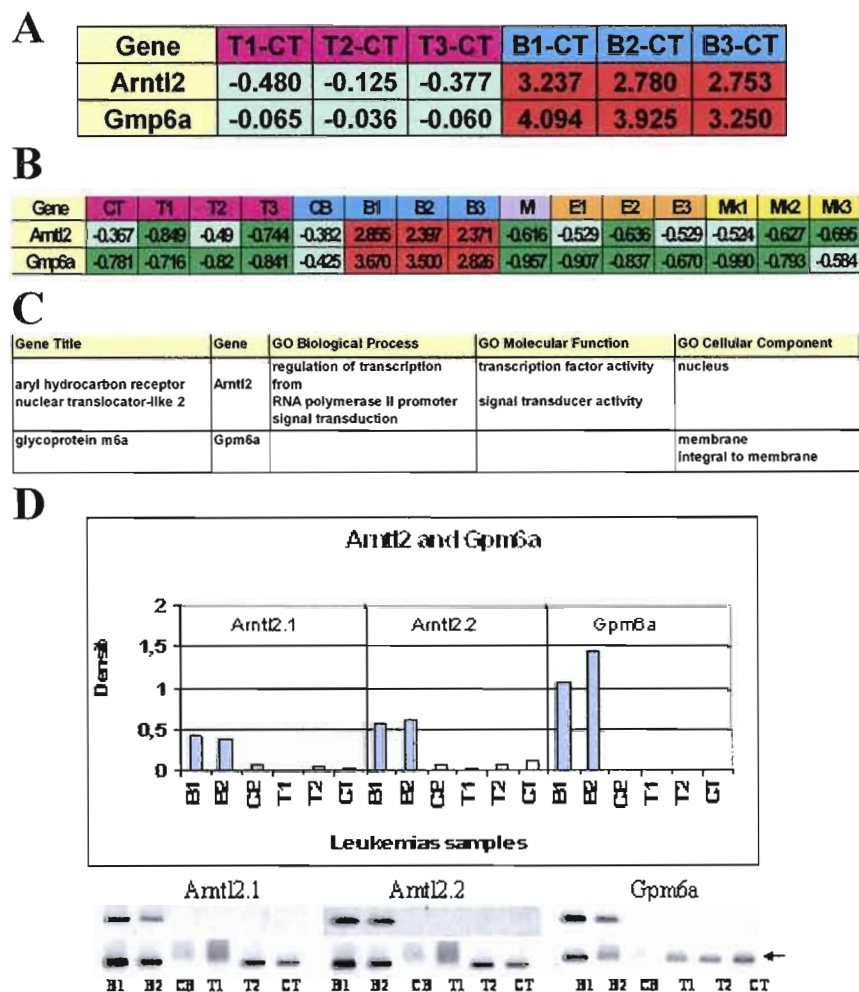


Figure 4.35: RT-PCR pour de 2 gènes spécifiques des leucémies B, Arntl2 et Gmp6a
 (A) Résultats RMA des ratios entre les leucémies B et le contrôle CB. (B) Résultats RMA des différents échantillons. Rouge: surexprimé; Vert: sous-exprimé. (C) Résultats de NETAFFY. (D) Amplification RT-PCR. Histogrammes correspondent à la quantification par rapport à β -actine et visualisation des bandes de RT-PCR (Cyndia Charfi).

4.7. Recherche de sites d'intégrations

Cette partie sur la recherche des sites d'intégration complète la partie des micropuces en essayant d'identifier les gènes qui sont responsables de l'apparition des leucémies mégacaryoblastiques. Le rétrovirus murin Graffi déclenche la leucémie parce qu'il s'intègre à proximité de proto-oncogènes (chapitre I, section 1.2.4). Les sites d'insertion rétrovirale dans les 3 leucémies mégacaryoblastiques analysées par micropuces ont été recherchés. Ainsi ces résultats reliés à ceux du profil d'expression génique pourraient permettre d'avoir une vue d'ensemble sur le processus qui a conduit à l'apparition de ce type particulier de leucémies.

4.7.1. Matériel et méthode

La méthode consiste à amplifier la région génomique bordant l'intégration rétrovirale (figure 4.36) et à localiser ainsi le locus d'insertion grâce à des navigateurs informatiques du génome de *Mus musculus* (EMSEMBL: www.ensembl.org and UCSC: www.genome.ucsc.edu)

La stratégie développée dans le cas du rétrovirus Graffi (projet de maîtrise de Diana Paulina Ospina Salazar, Laboratoire d'Éric Rassart) permet d'amplifier la région génomique bordant le côté 5' du rétrovirus. La technique de recherche des sites d'insertions a été utilisée telle que développée par Diana Paulina Ospina Salazar (adaptation de la technique d'Anton Berns). Le génome cellulaire est premièrement digéré par BamHI, une enzyme qui coupe fréquemment dans le génome de la souris, qui coupe dans le génome du rétrovirus Graffi mais ne coupe pas dans le LTR (figure 4.36 B). Une splinkerette est attachée par ligation à chaque extrémité cohésive laissée par la digestion de BamHI. La splinkerette est une amorce double-brin, qui a la caractéristique de comporter une structure en épingle à cheveux et qui correspond à une séquence qui n'existe pas dans le génome de la souris (elle provient du génome du virus de l'encéphalite et de l'arthrite caprine (CAEV)) (figure 4.36 A). Le fragment à amplifier par PCR (indiqué 'fragment d'intérêt' sur la figure 4.36 B) comprend une splinkerette, une région d'ADN génomique flanqué d'un 5'LTR. L'ADN est alors digéré par ClaI qui coupe dans le gène Pol ce qui va permettre d'éviter l'amplification par PCR du côté 3' du rétrovirus. Ensuite, une réaction PCR est réalisée grâce à une amorce correspondant à la

splinkerette et une autre placée dans la région U3 du virus Graffi. Seul le fragment d'intérêt est amplifié. Les fragments PCR sont alors clonés et séquencés et le positionnement sur le génome de *Mus musculus* est réalisé.

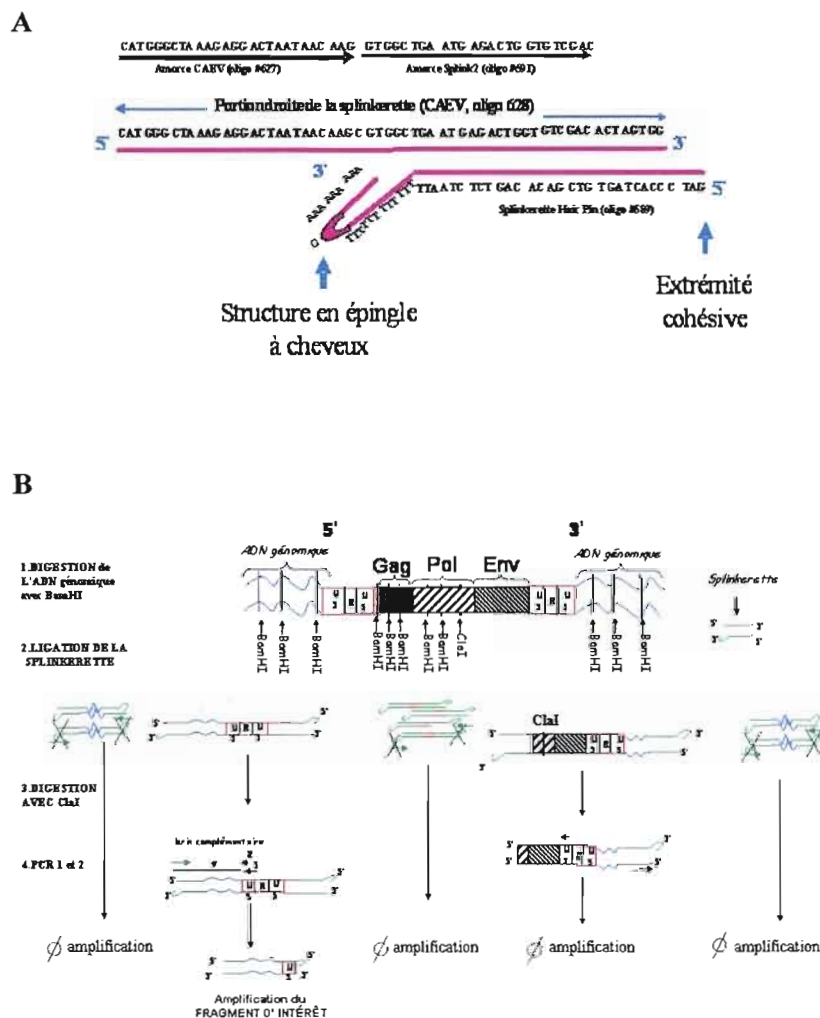


Figure 4.36: Recherche de sites d'intégrations. (A) Représentation de la splinkerette utilisée avec ses amorces respectives. (figure prise du mémoire de maîtrise de Diana Paulina Ospina Salazar et de Mikkers *et al*, 2002) (B) Processus d'amplification des régions génomiques flanquées d'une intégration rétrovirale (figure adaptée de la proposition de recherche de maîtrise de Diana Paulina Ospina Salazar).

4.7.2. Résultats

Onze intégrations dans la tumeur Mk1 (Kit⁺CD41⁺), 5 intégrations dans la tumeur Mk2 (Kit⁺CD41⁺) et 10 intégrations dans la tumeur Mk3 (Kit⁺CD41⁺) ont été clonées et séquencées. Le nombre élevé d'intégrations clonées par tumeur atteste du caractère polyclonal des tumeurs induites par le rétrovirus Graffi. Cette polyclonalité avait déjà été remarquée par analyse Southern Blot d'ADN de rates tumorales avec hybridation de la sonde U3 de Graffi (non montré). Plus d'une dizaine de bandes pouvaient être observées sans qu'elles soient présentes dans les échantillons contrôles. Certains fragments clonés correspondaient à l'amplification d'une séquence génomique et d'un virus endogène. Ces fragments ont été enlevés de l'analyse. Les tableaux 4.27, 4.28, 4.29 montrent les intégrations clonées avec leur localisation génomique pour les tumeurs Mk1, Mk2 et Mk3 respectivement.

Tableau 4.27: Intégrations rétrovirales dans Mk1

	Chromosome	gènes en amont*	dans le gène	gènes en aval*
1	1D	Gm817 (17kb), Lrrfip1(60kb)	Ramp1	Ube2f (95kb), Scly (130kb)
2	1G2	nfl2 (188kb), 1190005F20Rik (230k	Niban	Edem3 (92kb), Q8K3Do(199kb)
3	4B1	8H24Rik (133kb), 1700066J25Rik	Dnaic1	Cntfr (29kb), Dcnt3 (86kb)
4	5E 1	Pdgfra (216kb), Gsh2 (382kb)	-	Kit (114kb), Kdr (472kb)
5	6F3	Klrblb (10kb),Clec2h (159kb)	-	Clec21 (50kb), Klrblf (208kb)
6	6G1	Gsg1 (61kb), Hebp1 (130kb)	-	Pbp2(10 kb), Dynlt1 (50kb)
7	11B4	Pfn1(0.8kb), Rnf167(4kb)	-	Eno3(1.7kb), Spag7 (8.2kb)
8	11B5	Rhot1 (55kb), Rnf135(123kb)	Rhbd13	Zfp207(60kb), Psmd11(105kb)
9	15D3	Pvt1(211kb) Myc (463kb),	-	-
10	15F2	Dbx2(100kb), Nell2(225kb)	-	Tmem16f(36kb)
11	16B3	Heg1 (13kb), Slc12a8(118kb)	-	Muc13(12kb), Itgb5(48kb)

*:en amont ou en aval par rapport à la localisation chromosomique de l'insertion

en rouge: inclus dans RTCGD

Tableau 4.28: Intégrations rétrovirales dans Mk2

	Chromosome	gènes en amont*	dans le gène	gènes en aval*
1	1 E4	Pfkfb2(0.4kb), C4bp (54kb)	-	Yod1(0,8 kb). AA986890(1kb)
2	6A3	Pot 1a(14kb), Grp37(120kb)	-	-
3	8 E1	Irf8(81kb), Cox4il(149kb)	-	Foxfla(269kb). Mthfsd(280kb)
4	12C1	Hectd1(120kb), EG544864(35kb)	-	Heatr5a(11kb)6530401NO4Rik(124kb)
5	13D2.3	Pelo(46kb), Itgal(33kb)	-	-

*:en amont ou en aval par rapport à la localisation chromosomique de l'insertion
en rouge: inclus dans RTCGD

Tableau 4.29: Intégrations rétrovirales dans Mk3

	Chromosome	gènes en amont*	dans le gène	gènes en aval*
1	1C5	Glrip1(146kb), Spp2(230kb)	-	Arl14c(161kb), Sh3bp4(413kb)
2	1F	-	-	Ptprc(257kb), Atp6v1g3(466kb)
3	2H3	Prkebp1(43kb), Eya2(156kb)	-	Ncoa3 (65kb), Sulf2 (146kb)
4	5B1	Rheb(23kb), Cryng(83kb)	-	Prkag2(20kb)1500035N22Rik(144kb)
5	6D2	Rpn1(16kb), EG434064(115kb)	-	GATA-2 (72kb), Dnajb8(100kb)
6	7B4	Bax(21kb), Flt1(28kb)	Dhdh	Nucb1(4kb), Tulp2 (25kb)
7	7F5	Cend1(107kb), Oraov1(167kb)	-	Tpcn2(167kb), Mrgprf(253kb)
8	13A1	Calm13(27kb)	-	Calm4(6kb)Calm5(22kb)Net1(53kb)
9	14C3	Prmt5(20kb), Rem2(57kb)	-	D14Ert500e(4kb), Jub(31kb)
10	19A	Ankrd13d(18kb), Ssh3(32kb)	Adrbk1	Fbxl13(15kb)

*:en amont ou en aval par rapport à la localisation chromosomique de l'insertion
en rouge: inclus dans RTCGD

Ces tableaux présentent les gènes en amont et en aval des insertions. Certaines insertions ont eu lieu également dans des gènes. Certains gènes sont colorés en rouge (13 gènes). Il s'agit de gènes dont une ou plusieurs intégrations rétrovirales (sites communs d'intégration ou non) ont été trouvées par d'autres modèles rétroviraux et qui sont répertoriées dans la base de données des intégrations rétrovirales, RTCGD (Retrovirus Tagged Cancer Gene Database, <http://rtcgd.abcc.ncifcrf.gov/>). Neuf de ces 13 gènes sont ceux situés de façon la plus proche de l'intégration rétrovirale.

Le tableau suivant (tableau 4.30) montre les résultats des micropuces pour ces 13 gènes enregistrés dans RTCGD et proche d'une intégration rétrovirale dans les leucémies. Le tableau montre les résultats de RTCGD (nombre d'intégrations rétrovirales et sites communs d'intégration) également.

Tableau 4.30: Sites d'insertions et résultats des micropuces

	nom	T1	T2	T3	B1	B2	B3	M	E1	E2	E3	Mk1	Mk2	Mk3	IS	CIS
Mk1	Ramp1	1,5	0,0	1,8	0,2	-3,5	0,7	2,4	-1,6	-2,5	-1,9	0,8	1,7	0,4	6	+
Mk1	Kit	-2,2	-2,1	-2,2	-2,2	-2,3	1,2	2,8	-1,7	-1,6	-2,2	3,8	4,6	4,0	3	-
Mk1	Pfn1	0,0	-0,2	0,4	0,1	-0,2	0,0	-0,3	0,3	0,2	0,0	-0,1	-0,3	0,2	?	+
Mk1	Myc	0,0	-0,2	1,4	-0,6	0,1	-0,7	-0,4	0,1	0,3	0,7	0,0	-0,2	-0,5	75	+
Mk1	Itgb5	-1,0	-1,3	-1,2	-1,2	-0,9	-1,1	3,6	1,3	-0,1	-0,5	-0,8	0,2	3,0	3	-
Mk2	Irf8	-1,0	-2,6	-0,3	3,7	2,7	2,7	3,4	-1,7	-1,6	-3,1	-3,4	0,0	1,1	?	+
Mk2	Foxfla	-0,2	-0,1	0,2	0,0	0,0	-0,1	0,1	-0,3	0,0	0,1	0,3	0,0	0,0	2	-
Mk2	Itga1	-1,0	-1,1	-0,5	2,1	1,6	0,3	1,3	-0,6	-0,3	-0,8	-0,8	-0,3	0,0	1	-
Mk3	Prkcbp1	0,1	-0,4	-1,1	0,1	0,0	0,1	-0,2	1,3	-0,5	0,1	-0,2	-0,2	0,7	3	+
Mk3	Prkag2	1,1	1,1	1,2	-0,3	0,5	-0,5	0,8	-2,5	-2,6	-1,7	-0,1	3,7	-0,7	1	-
Mk3	Gata2	-2,3	-2,0	-2,3	-2,1	-2,4	-1,9	-1,6	0,2	2,4	1,1	3,6	4,0	3,4	1	-
Mk3	Ccnd1	1,2	-0,8	-0,7	-0,4	-0,6	1,2	0,2	-0,5	-0,5	-0,4	1,6	-0,1	-0,1	34	+
Mk3	Adrbk1	0,0	0,0	0,0	1,1	1,0	1,0	0,7	-0,9	-0,8	-0,7	-0,8	0,0	-0,5	2	-

IS: site d'intégration; CIS: site commun d'intégration

Parmi ces 13 gènes, certains sont associés à la signalisation cellulaire et à la prolifération (Myc, Ccnd1, Foxfla, Prkcbp1, Prkag2, Pfn1, Kit), alors que d'autres sont plus associés à la fonction hématopoïétique (Irf8, Itga1, Itgb5, Gata2, Kit). Ainsi on peut émettre l'hypothèse qu'ils peuvent se répartir selon les 2 groupes d'oncogènes: ceux impliqués dans la prolifération et la survie et l'autre dans l'immortalisation ou le blocage de la différenciation (chapitre I, section, 1.1.2.4). Ainsi, il est normal d'observer sur les résultats des micropuces que les gènes associés à la fonction hématopoïétique montrent des variations d'expressions significatives selon les leucémies alors que l'autre groupe en montre moins. Par exemple, on peut supposer que les expressions de Ccnd1 (Cyclin D1) et Myc sont très fortes dans toutes les leucémies mais ne varient pas d'un type de leucémie à l'autre. Gata2 et Kit sont les 2 gènes qui montrent une spécificité pour les leucémies mégacaryoblastiques et vont être analysés un peu plus en détail ultérieurement.

Les 2 gènes Ramp1 (receptor (calcitonin) activity modifying protein 1) et Adrbk1 (adrenergic receptor kinase, beta 1) ont une intégration intronique du rétrovirus Graffi. Le rapport sur ces intégrations rétrovirales dans RTCGD indique également qu'elles sont toutes introniques, dans un cas altérant le cadre de lecture (Ramp1) et dans l'autre cas conservant le cadre de lecture (Adrbk1).

Pfn1 (profilin 1) joue un rôle dans l'assemblage de l'actine, il est essentiel pour la survie cellulaire et la division lors du développement des souris et aurait également un rôle de suppresseur de tumeur (Witke et al, 2001, Susuki et al. 2006).

Foxf1 (forkhead box F1a) est un facteur de transcription, il est essentiel à la migration des cellules mésenchymateuses et il régule également l'expression de l'intégrine $\beta 3$ (CD61) (Malin et al. 2007).

Irf8 (Interferon regulatory factor 8) est un facteur de transcription et il est exprimé de façon prédominante dans les cellules hématopoïétiques. Il joue un rôle très important dans la différenciation myéloïde (granulocytes versus monocytes) et une absence de ce gène conduit au développement de leucémies chroniques et aiguës myéloïdes humaines (chapitre I, section 1.1.1.6.2.1.3). Il est également associé à un site commun d'intégration rétrovirale. Ce CIS aurait comme rôle d'activer l'expression de IRF8 et donc ce gène aurait un rôle différent dans la lignée myéloïde et lymphoïde (Ma et al. 2006).

Une étude récente montre l'implication de Itga1 (Integrin $\alpha 1$) dans la différenciation mégacaryoblastique. Cette étude apporte la preuve de la méthylation et de la répression de ITGA1 et du gène PELO situé dans un intron de ITGA1 lors des étapes précoces de la différenciation des mégacaryocytes humains (Cheli et al. 2007). Ainsi on peut émettre l'hypothèse que l'intégration du rétrovirus Graffi proche de ce locus pourrait empêcher la répression de ces 2 gènes, ce qui conduit en un blocage de la différenciation des mégacaryocytes. Itga2 fait partie de ce locus et est située à 200kb de l'intégration. Les souris *knock-out* pour Itga2 ont un défaut spécifique des plaquettes avec une mauvaise fixation aux fibres de collagène (Holtkotter et al. 2002). Cela montre que tout ce locus semble impliqué dans des fonctions de la lignée mégacaryocytaire.

Comme les tumeurs sont polyclonales, on peut penser que certaines intégrations (événement déclenchant du processus de cancer) sont présentes dans toutes les sous-populations et que d'autre sont présentes chacune dans des sous populations et ont aidé au processus de leucémogénèse. Le clonage de ces sites n'apporte pas la preuve de leur implication dans le processus de leucémogénèse et d'autres expérimentations doivent être effectuées dans ce but.

4.7.3. Vérification des sites d'intégrations 5E1 et 6D2

Les sites d'intégrations au niveau des chromosomes 8E1 et 5D2 sont à proximité des gènes Kit et Gata2 respectivement. Quelques expériences ont été effectuées pour confirmer ces sites à cause de l'implication connue de ces gènes dans la lignée

mégacaryocytaire. Rappelons que Gata2 est fortement exprimé dans les mégacaryocytes et il est impliqué dans les étapes précoces de la mégacaryopoïèse (chapitre I, section 1.1.1.6.2.2.4). Une mutation dans Gata1 a été trouvée responsable de l'apparition des leucémies mégacaryoblastiques dans les individus souffrant du syndrome de Down (chapitre I, section 1.1.2.4.2.4) et donc l'implication de Gata2 dans le développement des leucémies mégacaryoblastiques n'est pas du tout improbable. Kit est également caractéristique des progéniteurs hématopoïétiques et est un marqueur spécifique des progéniteurs mégacaryocytaires. Il est aussi classifié comme oncogène. Il code pour un récepteur tyrosine kinase qui agit positivement sur la prolifération cellulaire. Il est aussi surexprimé dans d'autres cancers (notamment cancers du sein) et des inhibiteurs des tyrosines kinases comme l'imatinib qui ciblent Kit sont utilisés comme agents thérapeutiques (Lengyel et al. 2007).

La figure 4.37 présente les loci d'intégrations pour 8E1 et 5D2 et montre que le virus Graffi est intégré en 5' des gènes Kit et Gata2 dans une orientation inverse (figure 4.37 A). Cette figure montre également la vérification de l'existence des sites d'intégrations 8E1 et 5D2 dans Mk1 et Mk3 respectivement (figure 4.37 B et C): afin de confirmer les sites d'intégrations, des PCRs ont été effectuées avec des oligonucléotides situés dans U3 et GAG du rétrovirus et dans la séquence génomique flanquant l'intégration. Le gel a été transféré sur membrane et les fragments PCR ont été hybridés avec une sonde U3 radioactive (figure 4.37 B). Des amplifications des intégrations correspondent bien à Mk1 pour 8E1 (Kit) et à Mk3 pour 5D2 (Gata2). Un essai PCR a été effectué pour détecter l'intégration au niveau du 8E1 à partir d'ADN de cellules triées de Mk1 envoyées aux micropuces (les autres expérimentations ayant été effectuées avec de l'ADN issus de la rate infiltrée non triée). Des fragments spécifiques et d'intensité supérieure aux essais sur la rate non triée ont été amplifiés (figure 4.37 C). Afin de savoir si les cellules triées ont également un patron polyclonal, un Southern Blot a été réalisé avec l'ADN de ces cellules triées et une hybridation avec U3 (figure 4.37 D). Six bandes sont visibles dont une d'intensité plus forte que les autres. Donc les cellules triées représentent bien une population polyclonale comportant une intégration proche de Kit. Il n'est pas possible par manque de matériel de continuer l'expérimentation pour savoir si l'intégration est majoritaire ou minoritaire dans ces populations. Pourtant, il est à noter que des techniques ont été développées pour répondre à ces questions: il s'agit par exemple de la PCR-LAM (Linear Mediated PCR) développée par l'équipe de Cristof von Kalle et qui nécessite très

peu de matériel au départ (50 à 100ng). Cette technique permet non seulement d'amplifier les intégrations rétrovirales mais aussi de déterminer la clonalité ou la polyclonalité d'une population, elle permet de détecter le clone dominant par une technique de PCR compétitive et quantitative (Schmidt et al. 2003). La figure 4.37.E est un Southern Blot effectué sur l'ADN des cellules triées E1, E2, et E3 et qui montre également la polyclonalité des populations érythroïdes leucémiques triées. Aucun essai n'a été effectué sur l'ADN de cellules triées de Mk2 et Mk3 par manque de matériel.

Des essais de RT-PCR ont été effectués sur les ARN des cellules triées de Mk1, Mk2 et Mk3 (échantillons envoyés aux micropuces) pour tester l'expression de Kit et Gata2 (figure 4.38). Les conditions étaient telles que décrites précédemment (section 4.6.2). Comme attendu d'après les résultats des micropuces, il n'y a pas de différences significatives entre les échantillons Mk1, Mk2 et Mk3. Cependant l'expression dans les leucémies est beaucoup plus forte que dans celle de l'échantillon non leucémique CD41⁺. Mais un passage en PCR en temps réel permettrait une conclusion plus sûre.

Deux hypothèses peuvent expliquer le fait que l'expression n'est pas plus forte dans Mk1 pour Kit et Mk3 pour Gata2: 1) les intégrations rétrovirales n'ont pas d'influence sur les gènes 2) les intégrations rétrovirales ont le rôle de bloquer la répression de ces gènes mais non d'augmenter encore leur niveau d'expression.

Il n'est pas rare qu'un gène avec une intégration rétrovirale à proximité n'ait pas une expression plus forte dans les tumeurs réarrangées. Cela ne signifie pas forcément que l'intégration n'a pas d'utilité. Par exemple, dans les tumeurs induites par le virus Graffi, aucune variation entre les échantillons réarrangés ou non-réarrangés au locus de Myc n'a été détectée car cet oncogène est de toute façon très fortement exprimé dans tous les échantillons leucémiques (Denicourt et al., 1999).

Même si beaucoup de preuves doivent être apportées par différentes expérimentations sur l'incidence des intégrations, l'hypothèse 2 est tout à fait probable. Kit et Gata2 sont des gènes fortement exprimés dans les progéniteurs mégacaryocytaires et leur expression diminue avec la différenciation. Un blocage de la répression de ces gènes peut favoriser le processus de la leucémogenèse, agissant sur l'immortalisation et/ou la prolifération des cellules leucémiques mégacaryoblastiques.

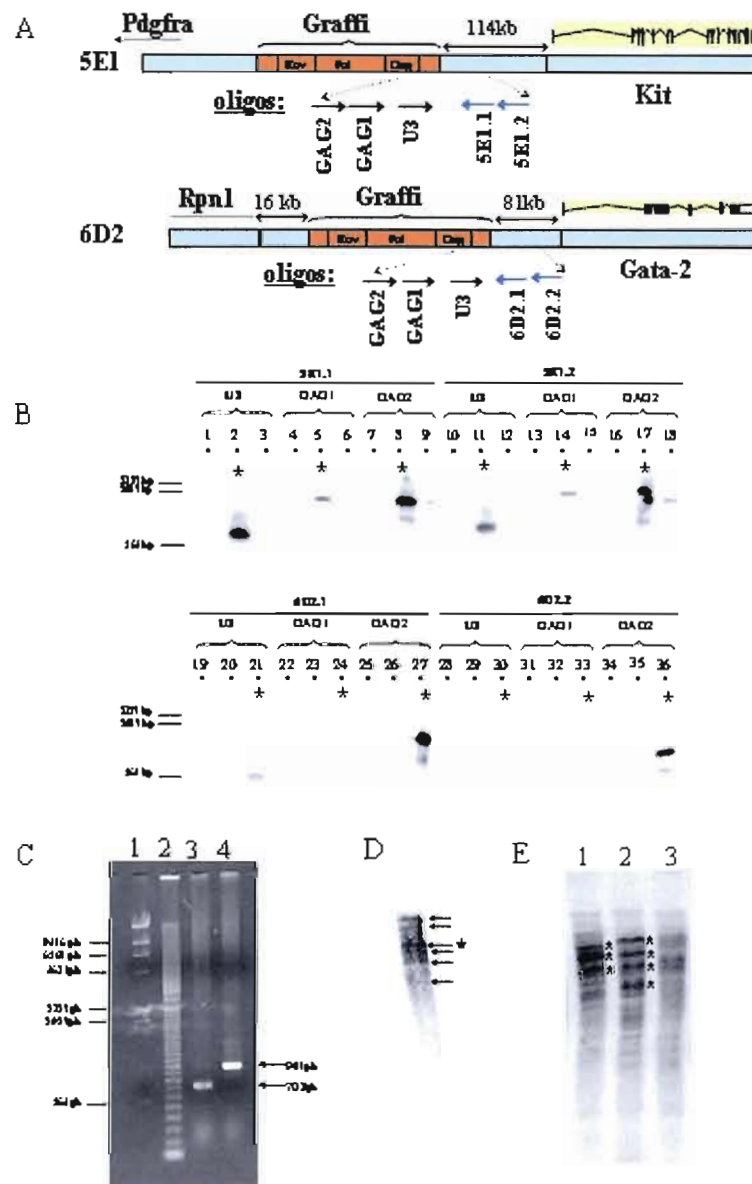


Figure 4.37: Sites d'intégrations 5E1 (Kit) et 6D2 (Gata2). (A) Présentation des 2 loci d'intégration. (B) Hybridation avec U3 des PCRs vérifiant les intégrations de Graffi dans les sites 5E1 et 6D2. Partie du haut: 5E1, partie du bas: 6D2. Mk1:2, 5, 8, 11, 14, 17, 20, 23, 26, 29, 32, 35; Mk2: 3, 6, 9, 12, 15, 18, 21, 24, 27, 31, 33, 36;; Mk3: 1, 4, 7, 10, 13, 16, 19, 22, 25, 28, 31, 34.(C) amplification PCR de 5E1 dans les cellules triées de Mk1. (D) Southern blot (hybridation avec U3) des cellules triées de Mk1. (E) Southern blot (hybridation avec U3) des cellules triées de E1, E2, E3).

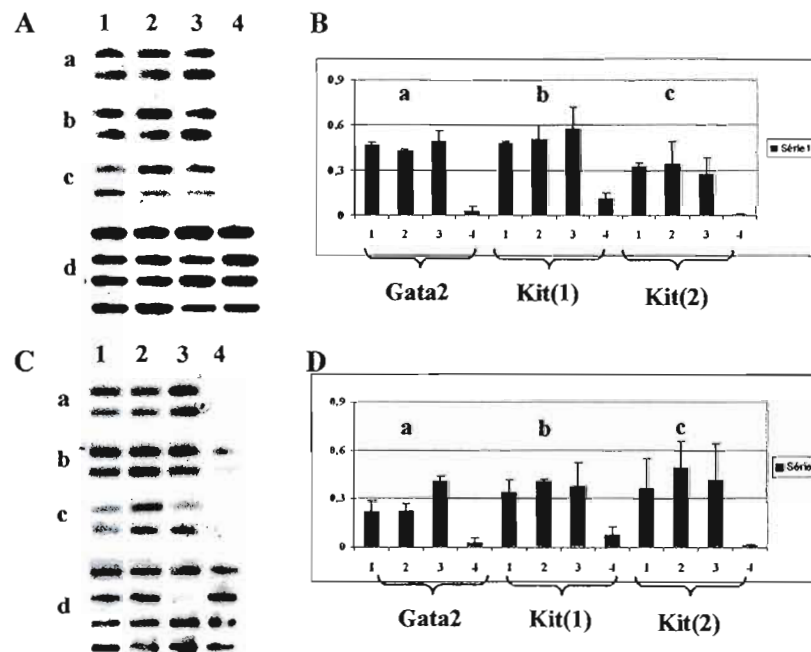


Figure 4.38: RT-PCR de Kit et Gata2 dans Mk1, Mk2 et Mk3. 1:Mk1, 2:Mk2, 3: Mk3, 4: cellules normales CD41⁺ (a) Gata2 (b) Kit, paire d'oligos 1; (c) Kit, paire d'oligos 2 (d) β-actine. (A) RT-PCR à 25 cycles. (B) Histogramme de quantification par rapport à β-actine. (C) RT-PCR à 20 cycles.(D) Histogramme de quantification par rapport à β-actine.

En conclusion, aucune preuve directe de l'identification des premiers intervenants responsables des leucémies Mk1, Mk2 et Mk3 n'a été apportée. Cependant le résultat du clonage des sites d'intégrations montre des pistes très intéressantes. Des expérimentations avec quelques-uns des gènes identifiés dans cette section (Gata2, Kit Irf8, Itag1) pourraient peut-être permettre de mieux comprendre les premières étapes de la leucémogénèse mégacaryoblastique, phénomène encore très peu élucidé.

4.8 Discussion et perspectives

4.8.1. Résumé

L'ensemble du travail présenté dans cette thèse repose sur l'utilisation du rétrovirus murin Graffi afin d'étudier le processus de leucémie. Les chapitres II et III relatent les expériences effectuées pour établir le modèle et le chapitre IV montre plus spécifiquement comment ce modèle a été utilisé afin de découvrir de nouveaux éléments concernant le phénomène de la leucémie.

Le rétrovirus murin Graffi représente maintenant un modèle bien établi pour étudier les différents types de leucémie malgré la complexité apportée par le nombre élevé de leucémies mixtes ou biphénotypiques et la polyclonalité des tumeurs.

Les résultats des micropuces ont apporté de nombreuses données grâce à la comparaison des différents types de leucémie entre eux. Aucune étude de profilage génique des leucémies n'a encore été publiée avec cette approche.

La multitude de données obtenues grâce aux micropuces et les diverses orientations qui peuvent être prises dans l'analyse font qu'un grand travail analytique préliminaire a dû être effectué (présenté dans ce chapitre) avant que des directions précises de recherche puissent être entreprises.

La caractérisation des gènes des leucémies lymphoïdes obtenus grâce aux micropuces est continuée par Cyndia Charfi en doctorat dans le laboratoire d'Eric Rassart, projet que l'étudiante a débuté en maîtrise.

Le projet de caractérisation des leucémies érythroïdes et mégacaryoblastiques est également prometteur. La suite du projet consiste à étudier la fonction de certains gènes d'intérêt afin d'établir leurs rôles dans l'érythropoïèse, la mégacaryopoïèse ou dans le développement de la leucémie. Les orientations possibles à la suite de ce projet étant nombreuses et le travail n'étant pas encore publié à ce jour, certaines perspectives précises pour poursuivre le projet sur les leucémies érythroïdes et mégacaryoblastiques vont être présentées ci-dessous.

4.8.2. Perspectives

Le problème majeur provenant du design de l'expérience des micropuces est qu'il n'est pas possible à priori de savoir, parmi les gènes spécifiques d'un type de leucémie, lesquels sont liés à la pathologie et lesquels sont liés à une fonction des cellules différenciées et non leucémiques. Même s'il s'avère que la fonction de ces gènes n'est pas oncogénique mais liée aux fonctions normales des cellules, il est tout de même très intéressant de les étudier. Ceci est particulièrement vrai pour la lignée mégacaryocytaire car la mégacaryopoïèse et les mégacaryocytes restent encore très mystérieux.

Un des points importants est de vérifier que ces gènes découverts grâce à un modèle murin remplissent un rôle comparable chez l'humain. En effet, la finalité du projet est d'apporter des éléments nouveaux pour mieux comprendre l'hématopoïèse et les leucémies humaines. Les gènes ont été sélectionnés au départ parce qu'ils ont une grande similarité avec leurs homologues humains. Il est cependant nécessaire de vérifier leur expression dans des échantillons leucémiques humains et de vérifier également qu'ils sont spécifiques du même type de leucémie.

Les premiers gènes dont la fonction doit être étudiée sont ceux qui ont été sélectionnés et validés en RT-PCR (section 4.6.3.1). D'après les résultats de RT-PCR, si d'autres gènes doivent être validés, un passage en PCR en temps réel sur un échantillonnage égal ou plus grand que celui utilisé précédemment semble préférable. Il est à noter que si l'expression spécifique des gènes candidats par rapport aux leucémies érythroïdes ou mégacaryoblastiques est assez clairement établie, la validation de l'expression par rapport aux cellules non leucémiques reste encore à améliorer. De meilleurs échantillons érythroblastiques doivent être prélevés de nouveau à partir de la moelle osseuse de souris (érythroblastes Ter119⁺). Pour les mégacaryoblastes ou mégacaryocytes normaux, la population splénique CD41⁺ doit être mieux caractérisée (tester plus de marqueurs en cytométrie de flux). De plus, de récents protocoles ont été développés pour isoler des populations pures de mégacaryoblastes à partir de la moelle osseuse de souris. Cela implique un tri des cellules premièrement par sélection négative pour éliminer les cellules Lin⁺, suivie d'une sélection positive avec CD41 (Dumon et al. 2006, Nakorn et al. 2003). L'équipe de Weissman a trouvé que la fraction CD9⁺CD41⁺FcR^{lo}c-kit⁺Sca-1⁺IL7R⁻Thy1.1⁺Lin⁻ de la moelle osseuse donne principalement des mégacaryocytes et des plaquettes en culture (Nakorn et al. 2003). De plus, l'injection de TPO à des souris permet

d'enrichir la moelle osseuse en mégacaryoblastes (Carow et al.2001). Afin de purifier correctement ces populations, la technologie avec aimant-billes de EasySep® utilisée pour le tri des cellules envoyées aux micropuces est probablement insuffisante.

Un autre point important est l'identification précise du transcrit alternatif correspondant à la surexpression spécifique observée. En effet le type de micropuces utilisé est performant et les *probes* sont sélectionnés de façon à être bien spécifiques du gène en question. Cependant, des incertitudes peuvent apparaître si plusieurs transcrits alternatifs existent. Affymetrix a d'ailleurs développé de nouvelles puces permettant de mieux différencier les transcrits. Ainsi différentes RT-PCR ou des techniques de RACE (Rapid Amplification of cDNA Ends) peuvent être appliquées pour identifier les transcrits en question. Il est important que le transcrit impliqué chez l'humain soit caractérisé.

Pour de nombreux gènes, la fonction est déjà connue à des degrés variables et cette connaissance permet d'établir des stratégies différentes selon les gènes.

Trois cas de figures générales peuvent être imaginés concernant la fonction des gènes sélectionnées: 1) ils peuvent avoir un rôle dans la prolifération, la survie des cellules. 2) ils peuvent avoir un rôle dans la différenciation. 3) leur rôle peut être associé à une fonction spécialisée des cellules matures.

4.8.2.1 Implication potentielle du gène dans la prolifération, la migration ou la survie des cellules

Cette étude permet de tester les capacités oncogéniques des gènes en question. Plusieurs essais ont été développés durant ces 20 dernières années pour caractériser les oncogènes. Il s'agit de tester les potentiels transformants et immortalisants des gènes, l'effet sur la migration, la survie, l'inhibition de contact et la croissance de cellules transfectées avec les gènes d'intérêt.

4.8.2.1.1. Tests de transformation cellulaire

Cet essai est effectué en co-transfectant le gène candidat avec Myc ou un gène Ras activé dans des fibroblastes embryonnaires murins. Si le gène a des propriétés

oncogéniques, des foyers de transformation vont apparaître attestant de l'acquisition de l'inhibition de contact et de la prolifération des cellules.

Ce test est global et mesure plusieurs paramètres à la fois. Mais il est possible de tester individuellement l'effet de la transfection du gène sur la prolifération cellulaire (comptage des cellules ou mesure de l'incorporation de thymidine radioactive ou de BrdU) ou sur l'apoptose (essai TUNEL ou analyses en cytométrie de flux). On peut également tester la capacité de migration des cellules transfectées sur de l'agar mou.

4.8.2.1.2. Tests sur la mobilité des cellules

Si l'hypothèse que le gène candidat puisse être impliqué dans la mobilité cellulaire (implication dans la régulation ou réorganisation du cytosquelette) semble probable et donc qu'il peut être impliqué dans le caractère invasif de la maladie, deux essais peuvent être réalisés: un essai de motilité avec la chambre de Boyden ou un essai d'invasion dans du Matrigel®.

La chambre de Boyden est composée de 2 espaces séparés d'une membrane. Des agents chémo-attractifs sont placés dans la chambre du bas attirant les cellules qui vont alors franchir la membrane. La capacité migratoire des cellules transfectées va être comparée à celle de cellules non-transfectées (Overholtzer et al. 2006). L'essai d'invasion du Matrigel est basé sur le même principe, avec une observation du nombre de cellules qui ont migré dans les sphéroïdes du gel (Giordano et al. 2000, Overholtzer et al. 2006).

4.8.2.2. Implication dans le processus de différenciation

Si le gène candidat n'est pas directement impliqué dans la prolifération cellulaire, il peut être impliqué dans la différenciation. Son rôle peut se situer dans les étapes précoces ou tardives de la différenciation. Cet aspect est très important car les approches thérapeutiques qui permettent la réactivation de la différenciation peuvent être très efficaces et peu toxiques. Connaître à quel moment est exprimé un gène lors de la différenciation peut également donner de grands indices sur sa fonction normale.

Plusieurs lignées cellulaires de différentes origines hématopoïétiques sont disponibles commercialement qui peuvent se différencier par addition de facteurs externes. Par exemple la lignée murine érythroblastique MEL peut réactiver sa différenciation en

érythrocytes par ajout de DMSO dans le milieu de culture (Tsiftoglou et al. 1985). Les lignées humaines K562 et HEL peuvent se différencier soit en suivant la différenciation érythroïde avec ajout de hémine (*hemin*) soit en suivant la différenciation mégacaryocytaire avec ajout de PMA (*phorbol myristate acetate*) (Huo et al. 2006).

De plus, d'autres lignées cellulaires commerciales qu'il serait intéressant d'utiliser existent. On peut citer TF1, AP214, LAMA-84 pour les lignées érythroïdes et CMK, Meg01, et SET2 pour les lignées mégacaryoblastiques. Toutes ces lignées dérivent de patients humains. La première étape est donc de tester l'expression des gènes candidats dans ces lignées humaines avant de commencer les tests de différenciation.

4.8.2.2.1 Expression des gènes candidats au niveau humain

L'expression des gènes candidats se fait par RT-PCR sur les ARN des lignées cellulaires. Il est important de ne pas se limiter aux leucémies érythroblastiques et mégacaryoblastiques et de tester l'expression de ces gènes dans des lignées humaines T, B et myéloïde. Ceci permet de reproduire les conditions utilisées pour les micropuces et de voir si les gènes candidats sont spécifiques des types érythroïde ou mégacaryoblastique chez l'humain également. Les lignées lymphoïdes humaines disponibles peuvent être Jurkat pour les lymphomes T, Daudi et Raji pour les leucémies B, U937 pour la lignée myéloïde.

Après cette expérience, il serait intéressant de pouvoir tester l'expression des gènes candidats sur des échantillons ARN de patients humains non passés en culture. Ces échantillons peuvent être acquis à partir de banques de tumeurs qui existent au Canada (www.ctrnet.ca) et notamment au Québec (La Banque de Cellules Leucémiques du Québec). Cependant il faut noter que les échantillons mégacaryoblastiques sont rares et donc difficiles à obtenir.

Il est aussi important d'étudier la corrélation entre la présence de molécules d'ARN et la production de protéines fonctionnelles. Ainsi, si des anticorps sont disponibles, il faudra faire des analyses Western-blot dans un premier temps pour vérifier la présence de la protéine.

4.8.2.2.2. Tests de différenciation

Si les gènes endogènes sont exprimés dans les lignées cellulaires et échantillons humains testés, il est intéressant d'étudier l'évolution de la courbe d'expression au cours de la différenciation.

Si au contraire, les gènes candidats ne sont pas exprimés dans les lignées cellulaires ni dans les échantillons testés, on peut tout de même émettre l'hypothèse qu'ils sont impliqués dans des étapes plus tardives de la différenciation ou dans la fonction normale des cellules matures chez l'humain.

Afin d'effectuer les tests de différenciation, les cellules sont mises en culture puis les agents de différenciation sont ajoutés au milieu de culture. En général quelques jours suffisent pour que la différenciation s'effectue. Il a été noté plus haut que K562 et HEL peuvent se différencier soit vers la lignée érythroïde (ajout d'hémine) soit vers la lignée mégacaryocytaire (ajout de PMA). Les lignées érythroïdes TF-1 (Kitamura et al. 1989), AP-217 (Mossuz et al. 1997) et LAMA84 (Drexler et al. 2004) peuvent elles aussi continuer quelques étapes de différenciation en érythrocytes lors de l'ajout d'hémine. De même, les cellules Meg01 (Deutsch et al. 2000), SET-2 (Uozumi et al. 2000), et CMK (obtenue d'un patient souffrant du syndrome de Down, Sato et al. 1989) peuvent aussi avancer dans la voie de différenciation mégacaryocytaire et même produire des plaquettes avec l'ajout de PMA.

La différenciation des cellules peut être observée grâce au changement de morphologie des cellules: augmentation de la taille des cellules et noyau polylobé lors de la différenciation en mégacaryocyte et diminution de la taille pour la formation des globules rouges. La différenciation des globules rouges peut être également suivie par l'observation de la synthèse de l'hémoglobine. Des tests plus poussés peuvent être effectués comme l'analyse en cytométrie de flux avec différents marqueurs ou des colorations spécifiques (coloration à la benzidine pour les érythroblastes et chloroacetate pour les mégacaryocytes).

Un pool de cellules doit être prélevé chaque jour du test de différenciation à des temps précis afin d'obtenir l'ARN nécessaire aux tests de RT-PCR.

Il faut tout de même noter que les lignées cellulaires citées ci-dessus proviennent en général de patients souffrant de leucémie myéloïde chronique et dont une crise blastique (après une phase de rémission) a pris le phénotype AML6 (leucémie érythroïde) ou AML7 (leucémie mégacaryoblastique). Ainsi ces lignées peuvent porter de façon

concomitante des marqueurs de la lignée myéloïde, érythroïde et mégacaryocytaire. De plus, la différenciation induite par les agents est souvent incomplète. Cependant, les systèmes de lignées cellulaires offrent l'avantage de permettre de ne pas manquer de matériel et de pouvoir les manipuler (transfections, etc...).

Un test de différenciation plus proche de ce qui se passe *in vivo* est la différenciation à partir de progéniteurs issus de la moelle osseuse ou de cordons ombilicaux de donneurs sains. On peut ainsi produire des érythroblastes et des mégacaryoblastes à partir des cellules CD34⁺ humaines. La production de mégacaryocytes et de plaquettes à partir de sang de cordon ombilical riche en HSCs semble avoir été particulièrement optimisée ces dernières années. Cela consiste à obtenir la fraction cellulaire CD34⁺ et les mettre en culture avec des cellules stromales et un cocktail de cytokines (ligand de Flt-3, SCF, TPO) puis de les cultiver sans cellules stromales avec un ajout de IL1 pour provoquer la différenciation en plaquettes (Matsunaga et al. 2006).

4.8.2.2.3 Expressions endogène et exogène des gènes candidats

Après les tests de différenciation, 2 cas de figure peuvent se présenter schématiquement: 1) le gène est exprimé fortement dans les progéniteurs ou lignées leucémiques et son expression diminue progressivement lors de l'induction de la différenciation; 2) l'expression du gène augmente progressivement lors de la différenciation.

Dans le premier cas (qui correspond au cas de Gata2 ou Kit), cela signifie que le gène est impliqué dans le processus de différenciation et un blocage de la répression de ce gène peut participer au processus de leucémogénèse. Il est alors intéressant de voir si la répression de ce gène (ARN anti-sens) dans les lignées leucémiques peut faciliter le processus de différenciation. De façon contraire, il est possible de surexprimer le gène dans ces lignées et voir si cela empêche l'induction de la différenciation. Il est à noter que la transfection classique de lignée cellulaire est souvent peu efficace et il est préférable de recourir à des infections lentivirales ou adénovirales.

Si les résultats montrent clairement que le gène candidat a une implication directe dans le blocage de la différenciation, il sera important de déterminer à quelle voie de

signalisation il est relié. En effet chaque gène de cette voie de signalisation peut être une cible potentielle pour la thérapie.

4.8.2.3. Implication dans un processus spécialisé des cellules différenciées

Si l'expression du gène augmente avec la différenciation, il est très probable que le gène candidat ait un rôle dans la fonction normale de l'hématopoïèse. Il peut aussi être impliqué dans les étapes tardives de la différenciation comme la formation des plaquettes. Ainsi un *knock-down* de ce gène plutôt qu'une surexpression pourrait permettre de révéler la fonction normale du gène. En effet, la répression de ce gène peut conduire à la formation de cellules matures anormales ou non fonctionnelles. Le gène doit être réprimé tout au long du processus de différenciation (4 à 5 jours en moyenne).

Si aucun phénotype n'est observable ou si l'expérience est techniquement impossible, des tests fonctionnels autres doivent être effectués et découlent de ce qui est connu sur le gène: domaines fonctionnels; localisation cellulaire, partenaires d'interaction et études déjà publiées.

Ces études sont très importantes: la formation de plaquettes anormales conduit à un problème de coagulation et la formation de globules rouges conduit à l'asphyxie de l'organisme. Beaucoup de maladies du sang proviennent de la fonction anormale d'une protéine essentielle des plaquettes ou des globules rouges. Il est donc intéressant de voir si la fonction du gène étudié peut être relié à une maladie du sang existante.

4.8.2.4. Études à long-terme

Si les tests de transformation et de différenciation montrent que la protéine codée par le gène candidat a un rôle oncogénique sans équivoque soit avec une fonction immortalisante soit avec une fonction proliférante, il faudra absolument étudier la capacité de transformation avec des méthodes *in vivo*.

Il est possible de faire des transfections stables dans des cellules en culture. Si les cellules forment des foyers de transformations *in vitro*, il est possible de les injecter à des

souris immunodéficientes de façon sous-cutanée avec du Matrigel® et d'observer la croissance de la tumeur *in vivo*.

De façon encore plus élaborée, il s'agit de créer des progéniteurs hématopoïétiques transgéniques. Les cellules vont être injectées à des souris immunodéficientes afin qu'elles repeuplent la moelle osseuse. L'apparition de symptômes leucémiques est alors observée.

Ces expérimentations sont délicates et leur succès dépend de challenges techniques, elles sont cependant réalisées avec succès dans de nombreux laboratoires notamment pour prouver le rôle oncogénique des protéines de fusion issues des translocations chromosomiques (Stubbs et al., 2007). Leur succès dépend aussi de la puissance de l'oncogène. Ainsi, les expérimentations sont destinées à être effectuées seulement si le rôle puissant de l'oncogène ainsi que sa fonction ont été déterminés de façon approfondie *in vitro*.

D'autre part, la preuve ultime qu'un gène candidat joue un rôle direct dans la différenciation est apportée par la génération de souris *knock-out* (*knock-out* complet ou conditionnel). De nombreux exemples des informations apportés par les *knock-out* ont été cités en chapitre I (section 1.1.1.6).

4.8.2.5. Cas d'un suppresseur de tumeur

Si le gène candidat a été sélectionné sur la base d'un suppresseur de tumeur (diminution significative de l'expression des gènes par rapport aux contrôles), la capacité du gène à réduire la prolifération cellulaire, à inhiber le cycle cellulaire ou à agir positivement sur l'apoptose va être étudiée.

4.8.2.6. Etablissement de lignées cellulaires à partir du modèle Graffi

Comme cela a déjà été évoqué dans la discussion du chapitre II, il serait très intéressant d'obtenir des lignées cellulaires à partir des cellules leucémiques érythroïdes et

mégacaryoblastiques induites par le rétrovirus Graffi. Le premier avantage est que cela permet de ne pas manquer de matériel et de faire différentes expérimentations, comme par exemple, arriver à faire repartir les cellules dans une voie de différenciation. Deuxièmement, cela permet d'éviter de travailler avec des populations polyclonales.

4.8.2.7. Gènes sélectionnés

Parmi les gènes à étudier, les gènes sélectionnés pour les leucémies érythroïde et mégacaryoblastique et confirmés en RT-PCR sont particulièrement intéressants. On peut orienter un peu les recherches grâce à ce qui est connu de leurs fonctions.

De façon générale, on peut s'attendre à ce que Fkbp9, Btbd14a et Cda soient plus fortement exprimés dans les phases précoces de la différenciation, comme également Gata1, Gata2, Kit. On peut penser que l'expression de l'ARN et de la protéine correspondante de Gulp1, Ncf2 Ltbp2 augmente avec la différenciation car ils semblent plus associés à la fonction normale du gène alors que les gènes cités précédemment ont un profil fonctionnel plus oncogénique (excepté Cda). Il n'est pas possible, avant les résultats des tests, de savoir si Snca, Slamf1, Rabgef1 ont un rôle oncogénique ou non.

De façon plus détaillée pour les gènes mégacaryocytaires: il est important 1) d'étudier l'interaction possible de Gulp1 avec Lpr1, 2) de voir l'effet de la liaison à la rapamycine à Fkbp9 et tenter d'élucider la voie de signalisation en aval de ce gène, 3) d'étudier l'effet de la stimulation du NO sur Gucyl1a3 (et Gucyl1b3) ainsi que de tenter d'élucider la voie de signalisation en aval 4) de voir si la voie de signalisation de Gucyl1a3 régule le fonctionnement de la NADPH oxydase.

Les gènes Irf8 et Itga1 identifiés et situés à proximité d'une intégration rétrovirale méritent également d'être étudiés, avec une implication potentielle dans le blocage de la différenciation des leucémies.

Il faut noter aussi qu'il serait intéressant de faire des tests afin de confirmer l'expression des gènes communs aux leucémies Mk, M et B dans les mégacaryocytes humains et les leucémies AMKL humaines. Il n'est pas possible de savoir si ces expressions sont seulement spécifiques aux leucémies mégacaryoblastiques induites par le virus Graffi et résulteraient peut-être de certaines intégrations rétrovirales. Mais de

façon intéressante, ces gènes semblent partager une fonction commune et n'ont pas encore été découverts dans les mégacaryocytes.

Concernant les gènes érythroïdes, 1) il est impératif de commencer par vérifier l'expression de *Slamf1* sur les échantillons de leucémies AML6 humains car il n'est pas connu pour être exprimé par les érythrocytes normaux. Il pourrait alors servir de cible thérapeutique 2) il est important de vérifier l'expression de *Cda* non seulement dans les leucémies humaines érythroïdes mais aussi AMKL (non connue) car *Cda* est responsable d'une mauvaise réponse à la thérapie. Pour les autres gènes, il faudrait analyser les domaines fonctionnels des protéines pour trouver des directions de recherche précises.

4.8.3. Conclusion

Les perspectives concernant le projet du profilage génique des leucémies induites par le rétrovirus murin Graffi ont le potentiel de fournir un travail à temps plein pour plus d'une personne et peut certainement conduire à des résultats intéressants qui feront l'objet de publications.

Seulement quelques gènes ont été validés par RT-PCR avec le but principal de valider l'expérience des micropuces. De nombreux gènes autres que ceux sélectionnés dans cette analyse, peuvent être un sujet d'étude. C'est pourquoi une attention particulière a été portée dans la rédaction du chapitre IV à montrer l'ensemble des données. L'ère post-génomique consiste à analyser la ou les fonctions des différents gènes codés par les génomes (annotation fonctionnelle du génome humain). L'avancement dans ce projet est observable en réanalysant les données des micropuces régulièrement. En effet, on s'aperçoit que des gènes avec une fonction inconnue quelques mois auparavant (gènes Riken) en ont subitement une bien définie quelques mois plus tard. Cependant, les protéines peuvent avoir des fonctions différentes selon le type cellulaire testé.

Les analyses des résultats des micropuces peuvent s'orienter dans plusieurs autres directions. Cela peut être une analyse plus précise de certaines voies de signalisation (avec observation de la phosphorylation de protéines et utilisation d'inhibiteurs enzymatique). Cela peut être également une étude des promoteurs des gènes spécifiques des leucémies érythroïdes ou mégacaryoblastiques. Cette étude se fait en observant la

présence de sites consensus sur les régions promotrices de tous les gènes spécifiques et peut se faire de façon automatisée.

4.9. Remerciements

Je tiens à remercier pour ce chapitre l'équipe du Dr. André Ponton du Centre Génomique McGill à Montréal pour leur excellent travail pour la réalisation des micropuces ainsi que Michal Blazejczyk et Mathieu Miron pour le développement du très performant logiciel FlexArray pour l'analyse des micropuces ainsi que pour leurs explications.

RÉFÉRENCES

- Abelson, H. T., et L. S. Rabstein. 1970. «Lymphosarcoma: virus-induced thymic-independent disease in mice». *Cancer Res.* vol. 30, no 8, p. 2213-2222.
- Abouzied, M. M., S. L. Baader, F. Dietz, J. Kappler, V. Gieselmann et S. Franken. 2004. «Expression patterns and different subcellular localization of the growth factors HDGF (hepatoma-derived growth factor) and HRP-3 (HDGF-related protein-3) suggest functions in addition to their mitogenic activity». *Biochem J.* vol. 378, no Pt 1, p. 169-176.
- Abraham, T. B., E. Freymuller, R. A. Mortara, J. D. Lopes et M. Mariano. 2003. «Morphological characterization of mouse B-1 cells». *Immunobiology.* vol. 208, no 4, p. 401-411.
- Affymetrix: Microarray Suite User, guide. 2001.
«<http://www.affymetrix.com/support/technical/manuals.affx>». vol. version 5.
- Aguilar, H. C., W. F. Anderson et P. M. Cannon. 2003. «Cytoplasmic tail of Moloney murine leukemia virus envelope protein influences the conformation of the extracellular domain: implications for mechanism of action of the R Peptide». *J Virol.* vol. 77, no 2, p. 1281-1291.
- Akashi, K., D. Traver, T. Miyamoto et I. L. Weissman. 2000. «A clonogenic common myeloid progenitor that gives rise to all myeloid lineages». *Nature.* vol. 404, no 6774, p. 193-197.
- Akyerli, C. B., M. Beksac, M. Holko, M. Frevel, K. Dalva, U. Ozbek, E. Soydan, M. Ozcan, G. Ozet, O. Ilhan, G. Gurman, H. Akan, B. R. Williams et T. Ozelik. 2005. «Expression of IFITM1 in chronic myeloid leukemia patients». *Leuk Res.* vol. 29, no 3, p. 283-286.
- Alarino, E. L., T. Kuukasjarvi, R. Karhu et A. Kallioniemi. 2007. «A comprehensive expression survey of bone morphogenetic proteins in breast cancer highlights the importance of BMP4 and BMP7». *Breast Cancer Res Treat.* vol. 103, no 2, p. 239-246.
- Albritton, L. M., L. Tseng, D. Scadden et J. M. Cunningham. 1989. «A putative murine ecotropic retrovirus receptor gene encodes a multiple membrane-spanning protein and confers susceptibility to virus infection». *Cell.* vol. 57, no 4, p. 659-666.
- Allman, D., J. Li et R. R. Hardy. 1999. «Commitment to the B lymphoid lineage occurs before DH-JH recombination». *J Exp Med.* vol. 189, no 4, p. 735-740.
- Andonegui, G., S. M. Kerfoot, K. McNagny, K. V. Ebbert, K. D. Patel et P. Kubes. 2005. «Platelets express functional Toll-like receptor-4». *Blood.* vol. 106, no 7, p. 2417-2423.

- Antonchuk, J., C. D. Hyland, D. J. Hilton et W. S. Alexander. 2004. «Synergistic effects on erythropoiesis, thrombopoiesis, and stem cell competitiveness in mice deficient in thrombopoietin and steel factor receptors». *Blood*. vol. 104, no 5, p. 1306-1313.
- Armstrong, S. A., et A. T. Look. 2005. «Molecular genetics of acute lymphoblastic leukemia». *J Clin Oncol*. vol. 23, no 26, p. 6306-6315.
- Ashman, L. K. 1999. «The biology of stem cell factor and its receptor C-kit». *Int J Biochem Cell Biol*. vol. 31, no 10, p. 1037-1051
- Athale, U. H., B. I. Razzouk, S. C. Raimondi, X. Tong, F. G. Behm, D. R. Head, D. K. Srivastava, J. E. Rubnitz, L. Bowman, C. H. Pui et R. C. Ribeiro. 2001. «Biology and outcome of childhood acute megakaryoblastic leukemia: a single institution's experience». *Blood*. vol. 97, no 12, p. 3727-3732.
- Avellino, R., S. Romano, R. Parasole, R. Bisogni, A. Lamberti, V. Poggi, S. Venuta et M. F. Romano. 2005. «Rapamycin stimulates apoptosis of childhood acute lymphoblastic leukemia cells». *Blood*. vol. 106, no 4, p. 1400-1406.
- Avraham, H., et D. J. Price. 1999. «Regulation of megakaryocytopoiesis and platelet production by tyrosine kinases and tyrosine phosphatases». *Methods*. vol. 17, no 3, p. 250-264.
- Azcoitia, V., M. Aracil, A. C. Martinez et M. Torres. 2005. «The homeodomain protein Meis1 is essential for definitive hematopoiesis and vascular patterning in the mouse embryo». *Dev Biol*. vol. 280, no 2, p. 307-320.
- Bahrami, S., M. Duch et F. S. Pedersen. 2004. «Change of tropism of SL3-2 murine leukemia virus, using random mutational libraries». *J Virol*. vol. 78, no 17, p. 9343-9351.
- Bakir, Hanaa Y., Chikako Tomiyama-Miyaji, Hisami Watanabe, Toru Nagura, Toshihiko Kawamura, Hiroho Sekikawa et Toru Abo (2006). Reasons why DBA/2 mice are resistant to malarial infection: expansion of CD3int B220+ γδ T cells with double-negative CD4 CD8 phenotype in the liver. 117: 127-135 p
- Baleeiro, R. B., P. C. Bergami-Santos, M. Y. Tomiyoshi, J. L. Gross, F. Haddad, C. A. Pinto, F. A. Soares, R. N. Younes et J. A. Barbuto. 2007. «Expression of a dendritic cell maturation marker CD83 on tumor cells from lung cancer patients and several human tumor cell lines: is there a biological meaning behind it?». *Cancer Immunol Immunother*.
- Barat, C., et E. Rassart. 1998. «Members of the GATA family of transcription factors bind to the U3 region of Cas-Br-E and gaffi retroviruses and transactivate their expression». *J Virol*. vol. 72, no 7, p. 5579-5588.
- Barat, C., et E. Rassart. 1998. «Nuclear factors that bind to the U3 region of two murine myeloid leukemia-inducing retroviruses, Cas-Br-E and Graffi». *Virology*. vol. 252, no 1, p. 82-95.

- Battinelli, E., S. R. Willoughby, T. Foxall, C. R. Valeri et J. Loscalzo. 2001. «Induction of platelet formation from megakaryocytoid cells by nitric oxide». *Proc Natl Acad Sci U S A*. vol. 98, no 25, p. 14458-14463.
- Becker, W., et H. G. Joost. 1999. «Structural and functional characteristics of Dyrk, a novel subfamily of protein kinases with dual specificity». *Prog Nucleic Acid Res Mol Biol*. vol. 62, p. 1-17.
- Begonja, A. J., L. Teichmann, J. Geiger, S. Gambaryan et U. Walter. 2006. «Platelet regulation by NO/cGMP signaling and NAD(P)H oxidase-generated ROS». *Blood Cells Mol Dis*. vol. 36, no 2, p. 166-170.
- Behrends, S., K. Vehse, H. Scholz, J. Bullerdiek et B. Kazmierczak. 2000. «Assignment of GUCY1A3, a candidate gene for hypertension, to human chromosome bands 4q31.1-->q31.2 by in situ hybridization». *Cytogenet Cell Genet*. vol. 88, no 3-4, p. 204-205.
- Bennett, J. M. 2000. «World Health Organization classification of the acute leukemias and myelodysplastic syndrome». *Int J Hematol*. vol. 72, no 2, p. 131-133.
- Bennett, J. M. 2002. «The FAB/MIC/WHO proposals for the classification of the chronic lymphoid leukemias». *Rev Clin Exp Hematol*. vol. 6, no 4, p. 330-334.
- Bennett, J. M., D. Catovsky, M. T. Daniel, G. Flandrin, D. A. Galton, H. R. Gralnick et C. Sultan. 1976. «Proposals for the classification of the acute leukaemias. French-American-British (FAB) co-operative group». *Br J Haematol*. vol. 33, no 4, p. 451-458.
- Bennett, J. M., D. Catovsky, M. T. Daniel, G. Flandrin, D. A. Galton, H. R. Gralnick et C. Sultan. 1985. «Proposed revised criteria for the classification of acute myeloid leukemia. A report of the French-American-British Cooperative Group». *Ann Intern Med*. vol. 103, no 4, p. 620-625.
- Bergeron, D., L. Poliquin, J. Houde, B. Barbeau et E. Rassart. 1992. «Analysis of proviruses integrated in Fli-1 and Evi-1 regions in Cas-Br-E MuLV-induced non-T-, non-B-cell leukemias». *Virology*. vol. 191, no 2, p. 661-669.
- Bertoni, A., S. Tadokoro, K. Eto, N. Pampori, L. V. Parise, G. C. White et S. J. Shattil. 2002. «Relationships between Rap1b, affinity modulation of integrin alpha IIb beta 3, and the actin cytoskeleton». *J Biol Chem*. vol. 277, no 28, p. 25715-25721.
- Bithell, A., J. Alberta, F. Hornby, C. D. Stiles et B. P. Williams. 2003. «Expression of the guanine nucleotide exchange factor, mr-gef, is regulated during the differentiation of specific subsets of telencephalic neurons». *Brain Res Dev Brain Res*. vol. 146, no 1-2, p. 107-118.
- Boag, J. M., A. H. Beesley, M. J. Firth, J. R. Freitas, J. Ford, K. Hoffmann, A. J. Cummings, N. H. de Klerk et U. R. Kees. 2006. «Altered glucose metabolism in childhood pre-B acute lymphoblastic leukaemia». *Leukemia*. vol. 20, no 10, p. 1731-1737.

- Bobkova, M., J. Stitz, M. Engelstadter, K. Cichutek et C. J. Buchholz. 2002. «Identification of R-peptides in envelope proteins of C-type retroviruses». *J Gen Virol.* vol. 83, no Pt 9, p. 2241-2246. >.
- Bond, H. M., M. Mesuraca, E. Carbone, P. Bonelli, V. Agosti, N. Amodio, G. De Rosa, M. Di Nicola, A. M. Gianni, M. A. Moore, A. Hata, M. Grieco, G. Morrone et S. Venuta. 2004. «Early hematopoietic zinc finger protein (EHZF), the human homolog to mouse Evi3, is highly expressed in primitive human hematopoietic cells». *Blood.* vol. 103, no 6, p. 2062-2070.
- Bonini, N. M., et B. I. Giasson. 2005. «Snaring the function of alpha-synuclein». *Cell.* vol. 123, no 3, p. 359-361
- Borrello, M. A., et R. P. Phipps. 1995. «Fibroblasts support outgrowth of splenocytes simultaneously expressing B lymphocyte and macrophage characteristics». *J Immunol.* vol. 155, no 9, p. 4155-4161.
- Bossolasco, M., F. Veillette, R. Bertrand et A. M. Mes-Masson. 2006. «Human TDE1, a TDE1/TMS family member, inhibits apoptosis in vitro and stimulates in vivo tumorigenesis». *Oncogene.* vol. 25, no 33, p. 4549-4558.
- Bour-Jordan, H., et J. A. Blueston. 2002. «CD28 function: a balance of costimulatory and regulatory signals». *J Clin Immunol.* vol. 22, no 1, p. 1-7.
- Brunning, R.D., E. Matutes et N.L. Harris. 2001. «World Health Organization Classification of Tunours». IARC Press. vol. 3, p. 77-80.
- Bundy, L. M., et H. Fan. 1998. «Molecular and phylogenetic analysis of SRS 19-6 murine leukemia virus». *Virus Genes.* vol. 18, no 1, p. 65-79.
- Bundy, L. M., M. Ru, B. F. Zheng, L. Cheng, P. K. Pattengale, J. L. Portis et H. Fan. 1995. «Biological characterization and molecular cloning of murine C-type retroviruses derived from the TSZ complex from mainland China». *Virology.* vol. 212, no 2, p. 367-382.
- Burgener, R., M. Wolf, T. Ganz et M. Baggiolini. 1990. «Purification and characterization of a major phosphatidylserine-binding phosphoprotein from human platelets». *Biochem J.* vol. 269, no 3, p. 729-734.
- Butcher, B. W., K. S. Wilson, S. H. Kroft, R. H. Collins, Jr. et V. Bhushan. 2005. «Acute leukemia with B-lymphoid and myeloid differentiation associated with an inv(5)(q13q33) in an adult patient». *Cancer Genet Cytogenet.* vol. 157, no 1, p. 62-66.
- Cailleau, R., et F. Costa. 1961. «Long-term in vitro cultivation of some mouse ascites tumors: Ehrlich ascites carcinoma». *J Natl Cancer Inst.* vol. 26, p. 271-281.
- Calattini, S., S. A. Chevalier, R. Duprez, S. Bassot, A. Froment, R. Mahieux et A. Gessain. 2005. «Discovery of a new human T-cell lymphotropic virus (HTLV-3) in Central Africa». *Retrovirology.* vol. 2, p. 30.

- Cantor, A. B., et S. H. Orkin. 2002. «Transcriptional regulation of erythropoiesis: an affair involving multiple partners». *Oncogene*. vol. 21, no 21, p. 3368-3376.
- Caron, A., J. F. Theoret, S. A. Mousa et Y. Merhi. 2002. «Anti-platelet effects of GPIIb/IIIa and P-selectin antagonism, platelet activation, and binding to neutrophils». *J Cardiovasc Pharmacol*. vol. 40, no 2, p. 296-306.
- Carow, C. E., N. E. Fox et K. Kaushansky. 2001. «Kinetics of endomitosis in primary murine megakaryocytes». *J Cell Physiol*. vol. 188, no 3, p. 291-303.
- Cerignoli, F., S. Rahmouni, Z. Ronai et T. Mustelin. 2006. «Regulation of MAP kinases by the VHR dual-specific phosphatase: implications for cell growth and differentiation». *Cell Cycle*. vol. 5, no 19, p. 2210-2215.
- Challier, C., L. Cocault, M. Flon, M. Pauchard, F. Porteu, S. Gisselbrecht et M. Souyri. 2000. «A new feature of Mpl receptor: ligand-induced transforming activity in FRE rat fibroblasts». *Oncogene*. vol. 19, no 16, p. 2033-2042.
- Chang, A. N., A. B. Cantor, Y. Fujiwara, M. B. Lodish, S. Droho, J. D. Crispino et S. H. Orkin. 2002. «GATA-factor dependence of the multitype zinc-finger protein FOG-1 for its essential role in megakaryopoiesis». *Proc Natl Acad Sci U S A*. vol. 99, no 14, p. 9237-9242.
- Chang, C. P., Y. Jacobs, T. Nakamura, N. A. Jenkins, N. G. Copeland et M. L. Cleary. 1997. «Meis proteins are major in vivo DNA binding partners for wild-type but not chimeric Pbx proteins». *Mol Cell Biol*. vol. 17, no 10, p. 5679-5687.
- Chatis, P. A., C. A. Holland, J. W. Hartley, W. P. Rowe et N. Hopkins. 1983. «Role for the 3' end of the genome in determining disease specificity of Friend and Moloney murine leukemia viruses». *Proc Natl Acad Sci U S A*. vol. 80, no 14, p. 4408-4411.
- Chatis, P. A., C. A. Holland, J. E. Silver, T. N. Frederickson, N. Hopkins et J. W. Hartley. 1984. «A 3' end fragment encompassing the transcriptional enhancers of nondefective Friend virus confers erythroleukemogenicity on Moloney leukemia virus». *J Virol*. vol. 52, no 1, p. 248-254.
- Chattopadhyay, S. K., M. R. Lander, E. Rands et D. R. Lowy. 1980. «Structure of endogenous murine leukemia virus DNA in mouse genomes». *Proc Natl Acad Sci U S A*. vol. 77, no 10, p. 5774-5778.
- Cheli, Y., S. Kanaji, B. Jacquelin, M. Chang, D. J. Nugent et T. J. Kunicki. 2007. «Transcriptional and epigenetic regulation of the integrin collagen receptor locus ITGA1-PELO-ITGA2». *Biochim Biophys Acta*. vol. 1769, no 9-10, p. 546-558.
- Cheng, J. D., M. Valianou, A. A. Canutescu, E. K. Jaffe, H. O. Lee, H. Wang, J. H. Lai, W. W. Bachovchin et L. M. Weiner. 2005. «Abrogation of fibroblast activation protein enzymatic activity attenuates tumor growth». *Mol Cancer Ther*. vol. 4, no 3, p. 351-360.
- Cheok, M. H., et W. E. Evans. 2006. «Acute lymphoblastic leukaemia: a model for the pharmacogenomics of cancer therapy». *Nat Rev Cancer*. vol. 6, no 2, p. 117-129.

- Chintala, S., J. Tan, R. Gautam, M. E. Rusiniak, X. Guo, W. Li, W. A. Gahl, M. Huizing, R. A. Spritz, S. Hutton, E. K. Novak et R. T. Swank. 2007. «The Slc35d3 gene, encoding an orphan nucleotide sugar transporter, regulates platelet-dense granules». *Blood*. vol. 109, no 4, p. 1533-1540.
- Chiocchetti, A., L. Gibello, A. Carando, A. Aspesi, P. Secco, E. Garelli, F. Loreni, M. Angelini, A. Biava, N. Dahl, U. Dianzani, U. Ramenghi, C. Santoro et I. Dianzani. 2005. «Interactions between RPS19, mutated in Diamond-Blackfan anemia, and the PIM-1 oncoprotein». *Haematologica*. vol. 90, no 11, p. 1453-1462.
- Chu, S., J. DeRisi, M. Eisen, J. Mulholland, D. Botstein, P. O. Brown et I. Herskowitz. 1998. «The transcriptional program of sporulation in budding yeast». *Science*. vol. 282, no 5389, p. 699-705.
- Chute, J. P., G. G. Muramoto, J. Whitesides, M. Colvin, R. Safi, N. J. Chao et D. P. McDonnell. 2006. «Inhibition of aldehyde dehydrogenase and retinoid signaling induces the expansion of human hematopoietic stem cells». *Proc Natl Acad Sci U S A*. vol. 103, no 31, p. 11707-11712.
- Clark, S. R., A. C. Ma, S. A. Tavener, B. McDonald, Z. Goodarzi, M. M. Kelly, K. D. Patel, S. Chakrabarti, E. McAvoy, G. D. Sinclair, E. M. Keys, E. Allen-Vercoe, R. Deviney, C. J. Doig, F. H. Green et P. Kubes. 2007. «Platelet TLR4 activates neutrophil extracellular traps to ensnare bacteria in septic blood». *Nat Med*. vol. 13, no 4, p. 463-469.
- Coffin, J.M., S.H. Hugues et H.E. Vamus. 1997. «Retroviruses». Cold Spring Harbor Laboratory Press.
- Cognasse, F., H. Hamzeh, P. Chavarin, S. Acquart, C. Genin et O. Garraud. 2005. «Evidence of Toll-like receptor molecules on human platelets». *Immunol Cell Biol*. vol. 83, no 2, p. 196-198.
- Cohen-Barak, O., D. T. Erickson, M. S. Badowski, D. A. Fuchs, C. L. Klassen, D. T. Harris et M. H. Brilliant. 2007. «Stem cell transplantation demonstrates that Sox6 represses epsilon y globin expression in definitive erythropoiesis of adult mice». *Exp Hematol*. vol. 35, no 3, p. 358-367.
- Combet, C., M. Jambon, G. Deleage et C. Geourjon. 2002. «Geno3D: automatic comparative molecular modelling of protein». *Bioinformatics*. vol. 18, no 1, p. 213-214. En ligne. <11836238>.
- Cope, L. M., R. A. Irizarry, H. A. Jaffee, Z. Wu et T. P. Speed. 2004. «A benchmark for Affymetrix GeneChip expression measures». *Bioinformatics*. vol. 20, no 3, p. 323-331.
- Cousins, D. J., T. H. Lee et D. Z. Staynov. 2002. «Cytokine coexpression during human Th1/Th2 cell differentiation: direct evidence for coordinated expression of Th2 cytokines». *J Immunol*. vol. 169, no 5, p. 2498-2506.

- Couve, A., S. Restituto, J. M. Brandon, K. J. Charles, H. Bawagan, K. B. Freeman, M. N. Pangalos, A. R. Calver et S. J. Moss. 2004. «Marlin-1, a novel RNA-binding protein associates with GABA receptors». *J Biol Chem.* vol. 279, no 14, p. 13934-13943.
- Cruz, R., L. Hedden, D. Boyer, M. G. Kharas, D. A. Fruman et K. K. Lee-Fruman. 2005. «S6 kinase 2 potentiates interleukin-3-driven cell proliferation». *J Leukoc Biol.* vol. 78, no 6, p. 1378-1385.
- Cumano, A., F. Dieterlen-Lievre et I. Godin. 1996. «Lymphoid potential, probed before circulation in mouse, is restricted to caudal intraembryonic splanchnopleura». *Cell.* vol. 86, no 6, p. 907-916.
- Cumano, A., et I. Godin. 2007. «Ontogeny of the hematopoietic system». *Annu Rev Immunol.* vol. 25, p. 745-785.
- de Bruijn, M. F., R. E. Ploemacher, A. E. Mayen, J. S. Voerman, W. A. Sliker, W. van Ewijk et P. J. Leenen. 1996. «High-level expression of the ER-MP58 antigen on mouse bone marrow hematopoietic progenitor cells marks commitment to the myeloid lineage». *Eur J Immunol.* vol. 26, no 12, p. 2850-2858.
- Debili, N., L. Coulombel, L. Croisille, A. Katz, J. Guichard, J. Breton-Gorius et W. Vainchenker. 1996. «Characterization of a bipotent erythro-megakaryocytic progenitor in human bone marrow». *Blood.* vol. 88, no 4, p. 1284-1296.
- DeCino, P., W. Lernhardt, H. Herbst et W. C. Raschke. 1988. «Expression of oncogenes in normal and transformed murine B lymphocytes». *Oncogene Res.* vol. 3, no 1, p. 33-37.
- Dedera, D. A., E. K. Waller, D. P. LeBrun, A. Sen-Majumdar, M. E. Stevens, G. S. Barsh et M. L. Cleary. 1993. «Chimeric homeobox gene E2A-PBX1 induces proliferation, apoptosis, and malignant lymphomas in transgenic mice». *Cell.* vol. 74, no 5, p. 833-843.
- Delabesse, E., S. Ogilvy, M. A. Chapman, S. G. Piltz, B. Gottgens et A. R. Green. 2005. «Transcriptional regulation of the SCL locus: identification of an enhancer that targets the primitive erythroid lineage in vivo». *Mol Cell Biol.* vol. 25, no 12, p. 5215-5225.
- Denicourt, C., E. Edouard et E. Rassart. 1999. «Oncogene activation in myeloid leukemias by Graffi murine leukemia virus proviral integration». *J Virol.* vol. 73, no 5, p. 4439-4442.
- Denicourt, C., C. A. Kozak et E. Rassart. 2003. «Gris1, a new common integration site in Graffi murine leukemia virus-induced leukemias: overexpression of a truncated cyclin D2 due to alternative splicing». *J Virol.* vol. 77, no 1, p. 37-44.
- Deutsch, V., M. Bitan, Y. Friedmann, A. Eldor et I. Vlodavsky. 2000. «Megakaryocyte maturation is associated with expression of the CXC chemokine connective tissue-activating peptide CTAP III». *Br J Haematol.* vol. 111, no 4, p. 1180-1189.

- DeVry, C. G., et S. Clarke. 1999. «Polymorphic forms of the protein L-isoaspartate (D-aspartate) O-methyltransferase involved in the repair of age-damaged proteins». *J Hum Genet.* vol. 44, no 5, p. 275-288.
- Di Pietro, S. M., et E. C. Dell'Angelica. 2005. «The cell biology of Hermansky-Pudlak syndrome: recent advances». *Traffic.* vol. 6, no 7, p. 525-533.
- Dialynas, D. P., Z. S. Quan, K. A. Wall, A. Pierres, J. Quintans, M. R. Loken, M. Pierres et F. W. Fitch. 1983. «Characterization of the murine T cell surface molecule, designated L3T4, identified by monoclonal antibody GK1.5: similarity of L3T4 to the human Leu-3/T4 molecule». *J Immunol.* vol. 131, no 5, p. 2445-2451
- Dickson, C., P. Acland, R. Smith, M. Dixon, R. Deed, D. MacAllan, W. Walther, F. Fuller-Pace, P. Kiefer et G. Peters. 1990. «Characterization of int-2: a member of the fibroblast growth factor family». *J Cell Sci Suppl.* vol. 13, p. 87-96.
- DiMartino, J. F., L. Sella, D. Traver, M. T. Firpo, J. Rhee, R. Warnke, S. O'Gorman, I. L. Weissman et M. L. Cleary. 2001. «The Hox cofactor and proto-oncogene Pbx1 is required for maintenance of definitive hematopoiesis in the fetal liver». *Blood.* vol. 98, no 3, p. 618-626..
- Doak, S. H., S. A. Jenkins, R. A. Hurle, M. Varma, A. Hawizy, H. G. Kynaston et J. M. Parry. 2007. «Bone morphogenic factor gene dosage abnormalities in prostatic intraepithelial neoplasia and prostate cancer». *Cancer Genet Cytogenet.* vol. 176, no 2, p. 161-165.
- Doll, J. A., V. M. Stellmach, N. P. Bouck, A. R. Bergh, C. Lee, L. P. Abramson, M. L. Cornwell, M. R. Pins, J. Borensztajn et S. E. Crawford. 2003. «Pigment epithelium-derived factor regulates the vasculature and mass of the prostate and pancreas». *Nat Med.* vol. 9, no 6, p. 774-780.
- Dong, Y., H. Zhang, A. C. Gao, J. R. Marshall et C. Ip. 2005. «Androgen receptor signaling intensity is a key factor in determining the sensitivity of prostate cancer cells to selenium inhibition of growth and cancer-specific biomarkers». *Mol Cancer Ther.* vol. 4, no 7, p. 1047-1055.
- Drexler, H. G., Y. Matsuo et R. A. MacLeod. 2004. «Malignant hematopoietic cell lines: in vitro models for the study of erythroleukemia». *Leuk Res.* vol. 28, no 12, p. 1243-1251..
- Dreyfus, F., B. Sola, S. Fichelson, P. Varlet, M. Charon, P. Tambourin, F. Wendling et S. Gisselbrecht. 1990. «Rearrangements of the Pim-1, c-myc, and p53 genes in Friend helper virus-induced mouse erythroleukemias». *Leukemia.* vol. 4, no 8, p. 590-594.
- Druker, B. J., M. Talpaz, D. J. Resta, B. Peng, E. Buchdunger, J. M. Ford, N. B. Lydon, H. Kantarjian, R. Capdeville, S. Ohno-Jones et C. L. Sawyers. 2001. «Efficacy and safety of a specific inhibitor of the BCR-ABL tyrosine kinase in chronic myeloid leukemia». *N Engl J Med.* vol. 344, no 14, p. 1031-1037

- Dumon, S., V. L. Heath, M. G. Tomlinson, B. Gottgens et J. Frampton. 2006. «Differentiation of murine committed megakaryocytic progenitors isolated by a novel strategy reveals the complexity of GATA and Ets factor involvement in megakaryocytopoiesis and an unexpected potential role for GATA-6». *Exp Hematol.* vol. 34, no 5, p. 654-663.
- Dunn, T. B. 1954. «Normal and pathologic anatomy of the reticular tissue in laboratory mice, with a classification and discussion of neoplasms». *J Natl Cancer Inst.* vol. 14, no 6, p. 1281-1433.
- Eisbacher, M., M. L. Holmes, A. Newton, P. J. Hogg, L. M. Khachigian, M. Crossley et B. H. Chong. 2003. «Protein-protein interaction between Fli-1 and GATA-1 mediates synergistic expression of megakaryocyte-specific genes through cooperative DNA binding». *Mol Cell Biol.* vol. 23, no 10, p. 3427-3441.
- Eisen, M. B., P. T. Spellman, P. O. Brown et D. Botstein. 1998. «Cluster analysis and display of genome-wide expression patterns». *Proc Natl Acad Sci U S A.* vol. 95, no 25, p. 14863-14868.
- Elagib, K. E., F. K. Racke, M. Mogass, R. Khetawat, L. L. Delehanty et A. N. Goldfarb. 2003. «RUNX1 and GATA-1 coexpression and cooperation in megakaryocytic differentiation». *Blood.* vol. 101, no 11, p. 4333-4341.
- Elzey, B. D., D. L. Sprague et T. L. Ratliff. 2005. «The emerging role of platelets in adaptive immunity». *Cell Immunol.* vol. 238, no 1, p. 1-9.
- Erickson, Loren D., Lorraine T. Tygrett, Sudershan K. Bhatia, Kenneth H. Grabstein et Thomas J. Waldschmidt (1996). Differential expression of CD22 (Lyb8) on murine B cells. 8: 1121-1129 p
- Erkeland, S. J., M. Valkhof, C. Heijmans-Antonissen, R. Delwel, P. J. Valk, M. H. Hermans et I. P. Touw. 2003. «The gene encoding the transcriptional regulator Yin Yang 1 (YY1) is a myeloid transforming gene interfering with neutrophilic differentiation». *Blood.* vol. 101, no 3, p. 1111-1117.
- Erkeland, S. J., M. Valkhof, C. Heijmans-Antonissen, A. van Hoven-Beijen, R. Delwel, M. H. Hermans et I. P. Touw. 2004. «Large-scale identification of disease genes involved in acute myeloid leukemia». *J Virol.* vol. 78, no 4, p. 1971-1980.
- Etzerodt, M., T. Mikkelsen, F. S. Pedersen, N. O. Kjeldgaard et P. Jorgensen. 1984. «The nucleotide sequence of the Akv murine leukemia virus genome». *Virology.* vol. 134, no 1, p. 196-207.
- Fass, D., R. A. Davey, C. A. Hamson, P. S. Kim, J. M. Cunningham et J. M. Berger. 1997. «Structure of a murine leukemia virus receptor-binding glycoprotein at 2.0 angstrom resolution». *Science.* vol. 277, no 5332, p. 1662-1666.
- Felsenstein, J. 1996. «Inferring phylogenies from protein sequences by parsimony, distance, and likelihood methods». *Methods Enzymol.* vol. 266, p. 418-427.

- Feuer, G., et P. L. Green. 2005. «Comparative biology of human T-cell lymphotropic virus type 1 (HTLV-1) and HTLV-2». *Oncogene*. vol. 24, no 39, p. 5996-6004.
- Fey, F. 1969. «Studies on the pathogenesis and mechanism of hematologic diversification by re-isolation of the myeloid leukemia virus (Graffi)». *Acta Haematol.* vol. 42, no 2, p. 65-75.
- Fey, F., et A. Graffi. 1965. «[Erythroblastic leukemia after the injection of mouse myelogenous leukemia virus]». *Z Krebsforsch.* vol. 67, no 2, p. 145-151.
- Fey, F., et A. Graffi. 1965. «[Research on Hematological Differentiation in Leukoses Induced by the Myeloid Leukemia Virus of Mice.]». *Acta Haematol.* vol. 33, p. 139-158.
- Fey, F., T. Schramm et A. Graffi. 1968. «[Hematologic diversification of induced leukemias following the in-vitro cultivation of the mouse myeloid leukemia virus]». *Arch Geschwulstforsch.* vol. 32, no 1, p. 4-10.
- Fiore-Donati, L., L. Chieco-Bianchi, G. Tridente et N. Pennelli. 1966. «Studies on thymus-dependent mechanisms of mouse leukemogenesis by Graffi virus». *Natl Cancer Inst Monogr.* vol. 22, p. 587-603.
- Fleming, T. J., M. L. Fleming et T. R. Malek. 1993. «Selective expression of Ly-6G on myeloid lineage cells in mouse bone marrow. RB6-8C5 mAb to granulocyte-differentiation antigen (Gr-1) detects members of the Ly-6 family». *J Immunol.* vol. 151, no 5, p. 2399-2408.
- Fredrickson, T. N., R. R. O'Neill, R. A. Rutledge, T. S. Theodore, M. A. Martin, S. K. Ruscetti, J. B. Austin et J. W. Hartley. 1987. «Biologic and molecular characterization of two newly isolated ras-containing murine leukemia viruses». *J Virol.* vol. 61, no 7, p. 2109-2119.
- Friend, C. 1957. «Cell-free transmission in adult Swiss mice of a disease having the character of a leukemia». *J Exp Med.* vol. 105, no 4, p. 307-318.
- Fuxa, M., et J. A. Skok. 2007. «Transcriptional regulation in early B cell development». *Curr Opin Immunol.* vol. 19, no 2, p. 129-136.
- Gantner, B. N., R. M. Simmons, S. J. Canavera, S. Akira et D. M. Underhill. 2003. «Collaborative induction of inflammatory responses by dectin-1 and Toll-like receptor 2». *J Exp Med.* vol. 197, no 9, p. 1107-1117.
- Gao, J., K. E. Zoller, M. H. Ginsberg, J. S. Brugge et S. J. Shattil. 1997. «Regulation of the pp72syk protein tyrosine kinase by platelet integrin alpha IIb beta 3». *Embo J.* vol. 16, no 21, p. 6414-6425.
- Gardner, M. B. 1978. «Type C viruses of wild mice: characterization and natural history of amphotropic, ecotropic, and xenotropic MuLV». *Curr Top Microbiol Immunol.* vol. 79, p. 215-259.

- Garin-Chesa, P., L. J. Old et W. J. Rettig. 1990. «Cell surface glycoprotein of reactive stromal fibroblasts as a potential antibody target in human epithelial cancers». *Proc Natl Acad Sci U S A*. vol. 87, no 18, p. 7235-7239.
- Garrett, S. C., K. M. Varney, D. J. Weber et A. R. Bresnick. 2006. «S100A4, a mediator of metastasis». *J Biol Chem*. vol. 281, no 2, p. 677-680.
- Gaur, A., et W. R. Green. 2003. «Analysis of the helper virus in murine retrovirus-induced immunodeficiency syndrome: evidence for immunoselection of the dominant and subdominant CTL epitopes of the BM5 ecotropic virus». *Viral Immunol*. vol. 16, no 2, p. 203-212.
- Ge, Y., M. L. Stout, D. A. Tatman, T. L. Jensen, S. Buck, R. L. Thomas, Y. Ravindranath, L. H. Matherly et J. W. Taub. 2005. «GATA1, cytidine deaminase, and the high cure rate of Down syndrome children with acute megakaryocytic leukemia». *J Natl Cancer Inst*. vol. 97, no 3, p. 226-231.
- Georgopoulos, K. 1997. «Transcription factors required for lymphoid lineage commitment». *Curr Opin Immunol*. vol. 9, no 2, p. 222-227.
- Gilliland, D. G., C. T. Jordan et C. A. Felix. 2004. «The molecular basis of leukemia». *Hematology Am Soc Hematol Educ Program*, p. 80-97.
- Giordano, S., A. Maffe, T. A. Williams, S. Artigiani, P. Gual, A. Bardelli, C. Basilico, P. Michieli et P. M. Comoglio. 2000. «Different point mutations in the met oncogene elicit distinct biological properties». *Faseb J*. vol. 14, no 2, p. 399-406.
- Godin, I., F. Dieterlen-Lievre et A. Cumano. 1995. «Emergence of multipotent hemopoietic cells in the yolk sac and paraaortic splanchnopleura in mouse embryos, beginning at 8.5 days postcoitus». *Proc Natl Acad Sci U S A*. vol. 92, no 3, p. 773-777.
- Godin, I. E., J. A. Garcia-Porrero, A. Coutinho, F. Dieterlen-Lievre et M. A. Marcos. 1993. «Para-aortic splanchnopleura from early mouse embryos contains B1a cell progenitors». *Nature*. vol. 364, no 6432, p. 67-70.
- Golemis, E. A., N. A. Speck et N. Hopkins. 1990. «Alignment of U3 region sequences of mammalian type C viruses: identification of highly conserved motifs and implications for enhancer design». *J Virol*. vol. 64, no 2, p. 534-542. .
- Golemis, E., Y. Li, T. N. Fredrickson, J. W. Hartley et N. Hopkins. 1989. «Distinct segments within the enhancer region collaborate to specify the type of leukemia induced by nondefective Friend and Moloney viruses». *J Virol*. vol. 63, no 1, p. 328-337.
- Golub, T. R., D. K. Slonim, P. Tamayo, C. Huard, M. Gaasenbeek, J. P. Mesirov, H. Coller, M. L. Loh, J. R. Downing, M. A. Caligiuri, C. D. Bloomfield et E. S. Lander. 1999. «Molecular classification of cancer: class discovery and class prediction by gene expression monitoring». *Science*. vol. 286, no 5439, p. 531-537.

- Gonzalez Garcia, J. R., S. K. Bohlander, M. Gutierrez Angulo, M. A. Esparza Flores, V. J. Picos Cardenas, J. P. Meza Espinoza, L. Ayala Madrigal Mde et H. Rivera. 2004. «A t(1;9)(q23.3 approximately q25;q34) affecting the ABL1 gene in a biphenotypic leukemia». *Cancer Genet Cytogenet.* vol. 152, no 1, p. 81-83
- Graf, B. A., D. A. Nazarenko, M. A. Borrello, L. J. Roberts, J. D. Morrow, J. Palis et R. P. Phipps. 1999. «Biphenotypic B/macrophage cells express COX-1 and up-regulate COX-2 expression and prostaglandin E(2) production in response to pro-inflammatory signals». *Eur J Immunol.* vol. 29, no 11, p. 3793-3803.
- Graf, T. 2002. «Differentiation plasticity of hematopoietic cells». *Blood.* vol. 99, no 9, p. 3089-3101.
- Graffi, A. 1957. «Chloroleukemia of mice». *Ann N Y Acad Sci.* vol. 68, no 2, p. 540-558.
- Graffi, A., F. Fey et T. Schramm. 1966. «Experiments on the hematologic diversification of viral mouse leukemias». *Natl Cancer Inst Monogr.* vol. 22, p. 21-31.
- Granger, S. W., L. M. Bundy et H. Fan. 1999. «Tandemization of a subregion of the enhancer sequences from SRS 19-6 murine leukemia virus associated with T-lymphoid but not other leukemias». *J Virol.* vol. 73, no 9, p. 7175-7184.
- Grasis, J. A., C. D. Browne et C. D. Tsoukas. 2003. «Inducible T cell tyrosine kinase regulates actin-dependent cytoskeletal events induced by the T cell antigen receptor». *J Immunol.* vol. 170, no 8, p. 3971-3976.
- Greene, L. M., W. O. Twal, M. J. Duffy, E. W. McDermott, A. D. Hill, N. J. O'Higgins, A. H. McCann, P. A. Dervan, W. S. Argraves et W. M. Gallagher. 2003. «Elevated expression and altered processing of fibulin-1 protein in human breast cancer». *Br J Cancer.* vol. 88, no 6, p. 871-878.
- Gross, L. 1951. «Pathogenic properties, and "vertical" transmission of the mouse leukemia agent». *Proc Soc Exp Biol Med.* vol. 78, no 1, p. 342-348.
- Gross, L. 1957. «Development and serial cellfree passage of a highly potent strain of mouse leukemia virus». *Proc Soc Exp Biol Med.* vol. 94, no 4, p. 767-771.
- Gu, Y. C., J. F. Talts, D. Gullberg, R. Timpl et M. Ekblom. 2001. «Glucocorticoids down-regulate the extracellular matrix proteins fibronectin, fibulin-1 and fibulin-2 in bone marrow stroma». *Eur J Haematol.* vol. 67, no 3, p. 176-184.
- Guibal, F. C., C. Moog-Lutz, P. Smolewski, Y. Di Gioia, Z. Darzynkiewicz, P. G. Lutz et Y. E. Cayre. 2002. «ASB-2 inhibits growth and promotes commitment in myeloid leukemia cells». *J Biol Chem.* vol. 277, no 1, p. 218-224.
- Guindon, S., F. Lethiec, P. Duroux et O. Gascuel. 2005. «PHYML Online—a web server for fast maximum likelihood-based phylogenetic inference». *Nucleic Acids Res.* vol. 33, no Web Server issue, p. W557-559.

- Guittet, O., P. Hakansson, N. Voevodskaya, S. Fridd, A. Graslund, H. Arakawa, Y. Nakamura et L. Thelander. 2001. «Mammalian p53R2 protein forms an active ribonucleotide reductase in vitro with the R1 protein, which is expressed both in resting cells in response to DNA damage and in proliferating cells». *J Biol Chem.* vol. 276, no 44, p. 40647-40651.
- Gurney, A. L., K. Carver-Moore, F. J. de Sauvage et M. W. Moore. 1994. «Thrombocytopenia in c-mpl-deficient mice». *Science.* vol. 265, no 5177, p. 1445-1447.
- Gutzmer, R., B. Kother, J. Zwirner, D. Dijkstra, R. Purwar, M. Wittmann et T. Werfel. 2006. «Human plasmacytoid dendritic cells express receptors for anaphylatoxins C3a and C5a and are chemoattracted to C3a and C5a». *J Invest Dermatol.* vol. 126, no 11, p. 2422-2429.
- Hall, M. A., D. J. Curtis, D. Metcalf, A. G. Elefanty, K. Sourris, L. Robb, J. R. Gothert, S. M. Jane et C. G. Begley. 2003. «The critical regulator of embryonic hematopoiesis, SCL, is vital in the adult for megakaryopoiesis, erythropoiesis, and lineage choice in CFU-S12». *Proc Natl Acad Sci U S A.* vol. 100, no 3, p. 992-997.
- Hanecak, R., P. K. Pattengale et H. Fan. 1988. «Addition of substitution of simian virus 40 enhancer sequences into the Moloney murine leukemia virus (M-MuLV) long terminal repeat yields infectious M-MuLV with altered biological properties». *J Virol.* vol. 62, no 7, p. 2427-2436.
- Hannuksela, J., S. Parkkila, A. Waheed, R. S. Britton, R. E. Fleming, B. R. Bacon et W. S. Sly. 2003. «Human platelets express hemochromatosis protein (HFE) and transferrin receptor 2». *Eur J Haematol.* vol. 70, no 4, p. 201-206.
- Hao, X., M. S. Shin, J. X. Zhou, C. H. Lee, C. F. Qi, Z. Naghashfar, J. W. Hartley, T. N. Fredrickson, J. M. Ward et H. C. Morse, 3rd. 2006. «Histologic and molecular characterizations of megakaryocytic leukemia in mice». *Leuk Res.* vol. 30, no 4, p. 397-406.
- Hardy, R. R., Y. S. Li, D. Allman, M. Asano, M. Gui et K. Hayakawa. 2000. «B-cell commitment, development and selection». *Immunol Rev.* vol. 175, p. 23-32.
- Hart, A., F. Melet, P. Grossfeld, K. Chien, C. Jones, A. Tunnacliffe, R. Favier et A. Bernstein. 2000. «Fli-1 is required for murine vascular and megakaryocytic development and is hemizygously deleted in patients with thrombocytopenia». *Immunity.* vol. 13, no 2, p. 167-177.
- Hartenstein, V. 2006. «Blood cells and blood cell development in the animal kingdom». *Annu Rev Cell Dev Biol.* vol. 22, p. 677-712.
- Hasegawa, S., R. Pawankar, K. Suzuki, T. Nakahata, S. Furukawa, K. Okumura et C. Ra. 1999. «Functional expression of the high affinity receptor for IgE (FcepsilonRI) in human platelets and its' intracellular expression in human megakaryocytes». *Blood.* vol. 93, no 8, p. 2543-2551.

- Hauser, W., K. P. Knobloch, M. Eigenthaler, S. Gambaryan, V. Krenn, J. Geiger, M. Glazova, E. Rohde, I. Horak, U. Walter et M. Zimmer. 1999. «Megakaryocyte hyperplasia and enhanced agonist-induced platelet activation in vasodilator-stimulated phosphoprotein knockout mice». *Proc Natl Acad Sci U S A*. vol. 96, no 14, p. 8120-8125.
- Henderson, J. R., D. Brown, J. A. Richardson, E. N. Olson et M. C. Beckerle. 2002. «Expression of the gene encoding the LIM protein CRP2: a developmental profile». *J Histochem Cytochem*. vol. 50, no 1, p. 107-111.
- Hentges, K. E., K. C. Weiser, T. Schountz, L. S. Woodward, H. C. Morse et M. J. Justice. 2005. «Evi3, a zinc-finger protein related to EBFAZ, regulates EBF activity in B-cell leukemia». *Oncogene*. vol. 24, no 7, p. 1220-1230.
- Hernandez-Hoyos, G., M. K. Anderson, C. Wang, E. V. Rothenberg et J. Alberola-Ila. 2003. «GATA-3 expression is controlled by TCR signals and regulates CD4/CD8 differentiation». *Immunity*. vol. 19, no 1, p. 83-94.
- Herr, W. 1984. «Nucleotide sequence of AKV murine leukemia virus». *J Virol*. vol. 49, no 2, p. 471-478.
- Herr, W., A. P. Perlmutter et W. Gilbert. 1983. «Monoclonal AKR/J thymic leukemias contain multiple JH immunoglobulin gene rearrangements». *Proc Natl Acad Sci U S A*. vol. 80, no 24, p. 7433-7436.
- Hirvonen, H., V. Hukkanen, T. T. Salmi, T. P. Makela, T. T. Pelliniemi, S. Knuutila et R. Alitalo. 1991. «Expression of L-myc and N-myc proto-oncogenes in human leukemias and leukemia cell lines». *Blood*. vol. 78, no 11, p. 3012-3020.
- Hitzler, J. K., et A. Zipursky. 2005. «Origins of leukaemia in children with Down syndrome». *Nat Rev Cancer*. vol. 5, no 1, p. 11-20.
- Ho, M. K., et T. A. Springer. 1982. «Mac-1 antigen: quantitative expression in macrophage populations and tissues, and immunofluorescent localization in spleen». *J Immunol*. vol. 128, no 5, p. 2281-2286.
- Hoffman, P. M., E. F. Cimino, D. S. Robbins, R. D. Broadwell, J. M. Powers et S. K. Russetti. 1992. «Cellular tropism and localization in the rodent nervous system of a neuropathogenic variant of Friend murine leukemia virus». *Lab Invest*. vol. 67, no 3, p. 314-321.
- Hogan, B. L. 1996. «Bone morphogenetic proteins: multifunctional regulators of vertebrate development». *Genes Dev*. vol. 10, no 13, p. 1580-1594.
- Holmes, K. L., W. Y. Langdon, T. N. Fredrickson, R. L. Coffman, P. M. Hoffman, J. W. Hartley et H. C. Morse, 3rd. 1986. «Analysis of neoplasms induced by Cas-Br-M MuLV tumor extracts». *J Immunol*. vol. 137, no 2, p. 679-688.
- Holmes, M. L., N. Bartle, M. Eisbacher et B. H. Chong. 2002. «Cloning and analysis of the thrombopoietin-induced megakaryocyte-specific glycoprotein VI promoter and its regulation by GATA-1, Fli-1, and Sp1». *J Biol Chem*. vol. 277, no 50, p. 48333-48341.

Holtkotter, O., B. Nieswandt, N. Smyth, W. Muller, M. Hafner, V. Schulte, T. Krieg et B. Eckes. 2002. «Integrin alpha 2-deficient mice develop normally, are fertile, but display partially defective platelet interaction with collagen». *J Biol Chem.* vol. 277, no 13, p. 10789-10794.

Hosen, N., C. Y. Park, N. Tatsumi, Y. Oji, H. Sugiyama, M. Gramatzki, A. M. Krensky et I. L. Weissman. 2007. «CD96 is a leukemic stem cell-specific marker in human acute myeloid leukemia». *Proc Natl Acad Sci U S A.* vol. 104, no 26, p. 11008-11013.

Hu, L., J. S. Sham, D. Xie, J. M. Wen, W. S. Wang, Y. Wang et X. Y. Guan. 2007. «Up-regulation of fibroblast growth factor 3 is associated with tumor metastasis and recurrence in human hepatocellular carcinoma». *Cancer Lett.*

Huo, X. F., J. Yu, H. Peng, Z. W. Du, X. L. Liu, Y. N. Ma, X. Zhang, Y. Zhang, H. L. Zhao et J. W. Zhang. 2006. «Differential expression changes in K562 cells during the hemin-induced erythroid differentiation and the phorbol myristate acetate (PMA)-induced megakaryocytic differentiation». *Mol Cell Biochem.* vol. 292, no 1-2, p. 155-167.

Ichikawa, M., T. Asai, T. Saito, S. Seo, I. Yamazaki, T. Yamagata, K. Mitani, S. Chiba, S. Ogawa, M. Kurokawa et H. Hirai. 2004. «AML-1 is required for megakaryocytic maturation and lymphocytic differentiation, but not for maintenance of hematopoietic stem cells in adult hematopoiesis». *Nat Med.* vol. 10, no 3, p. 299-304.

Ikeda, H., K. Kato, H. Kitani, T. Suzuki, T. Yoshida, Y. Inaguma, N. Yamamoto, J. G. Suh, B. H. Hyun, T. Yamagata, T. Namikawa et T. Tomita. 2001. «Virological properties and nucleotide sequences of Cas-E-type endogenous ecotropic murine leukemia viruses in South Asian wild mice, *Mus musculus castaneus*». *J Virol.* vol. 75, no 11, p. 5049-5058.

Ikeda, M., O. Ishida, T. Hinoi, S. Kishida et A. Kikuchi. 1998. «Identification and characterization of a novel protein interacting with Ral-binding protein 1, a putative effector protein of Ral». *J Biol Chem.* vol. 273, no 2, p. 814-821.

Ikonomi, P., C. E. Rivera, M. Riordan, G. Washington, A. N. Schechter et C. T. Noguchi. 2000. «Overexpression of GATA-2 inhibits erythroid and promotes megakaryocyte differentiation». *Exp Hematol.* vol. 28, no 12, p. 1423-1431.

Irizarry, R. A., B. M. Bolstad, F. Collin, L. M. Cope, B. Hobbs et T. P. Speed. 2003. «Summaries of Affymetrix GeneChip probe level data». *Nucleic Acids Res.* vol. 31, no 4, p. e15.

Ishida, M., T. Ishida, N. Ono, H. Matsuura, M. Watanabe, G. Kajiyama, M. Kambe et T. Oshima. 1996. «Effects of insulin on calcium metabolism and platelet aggregation». *Hypertension.* vol. 28, no 2, p. 209-212.

Ishikawa-Sekigami, T., Y. Kaneko, Y. Saito, Y. Murata, H. Okazawa, H. Ohnishi, P. A. Oldenborg, Y. Nojima et T. Matozaki. 2006. «Enhanced phagocytosis of CD47-deficient red blood cells by splenic macrophages requires SHPS-1». *Biochem Biophys Res Commun.* vol. 343, no 4, p. 1197-1200.

Izumoto, Y., T. Kuroda, H. Harada, T. Kishimoto et H. Nakamura. 1997. «Hepatoma-derived growth factor belongs to a gene family in mice showing significant homology in the amino terminus». *Biochem Biophys Res Commun.* vol. 238, no 1, p. 26-32.

Jahid, S., L. M. Bundy, S. W. Granger et H. Fan. 2006. «Chimeras between SRS and Moloney murine leukemia viruses reveal novel determinants in disease specificity and MCF recombinant formation». *Virology.*

Janeway, C., P. Travers, M. Walport et M. Schlochik. 2004. «Immunobiology, 6th edition». Garland Science.

Jeffery, I. B., D. G. Higgins et A. C. Culhane. 2006. «Comparison and evaluation of methods for generating differentially expressed gene lists from microarray data». *BMC Bioinformatics.* vol. 7, p. 359.

Jensen, G. S., A. R. Belch, F. Kherani, M. J. Mant, B. A. Ruether et L. M. Pilarski. 1992. «Restricted expression of immunoglobulin light chain mRNA and of the adhesion molecule CD11b on circulating monoclonal B lineage cells in peripheral blood of myeloma patients». *Scand J Immunol.* vol. 36, no 6, p. 843-853.

Jin, L., K. J. Hope, Q. Zhai, F. Smadja-Joffe et J. E. Dick. 2006. «Targeting of CD44 eradicates human acute myeloid leukemic stem cells». *Nat Med.* vol. 12, no 10, p. 1167-1174.

Jing, S., Y. Yu, M. Fang, Z. Hu, P. L. Holst, T. Boone, J. Delaney, H. Schultz, R. Zhou et G. M. Fox. 1997. «GFRalpha-2 and GFRalpha-3 are two new receptors for ligands of the GDNF family». *J Biol Chem.* vol. 272, no 52, p. 33111-33117.

Jo, D., M. S. Lyu, E. G. Cho, D. Park, C. A. Kozak et M. G. Kim. 2001. «Identification and genetic mapping of the mouse Fkbp9 gene encoding a new member of FK506-binding protein family». *Mol Cells.* vol. 12, no 2, p. 272-275.

Kaplan, H. S. 1964. «The Role of Radiation on Experimental Leukemogenesis». *Natl Cancer Inst Monogr.* vol. 14, p. 207-220.

Katoh, M. 2002. «Molecular cloning and characterization of ST7R (ST7-like, ST7L) on human chromosome 1p13, a novel gene homologous to tumor suppressor gene ST7 on human chromosome 7q31». *Int J Oncol.* vol. 20, no 6, p. 1247-1253.

Kaushansky, K. 2006. «Lineage-specific hematopoietic growth factors». *N Engl J Med.* vol. 354, no 19, p. 2034-2045.

Kawabata, H., R. S. Germain, T. Ikezoe, X. Tong, E. M. Green, A. F. Gombart et H. P. Koeffler. 2001. «Regulation of expression of murine transferrin receptor 2». *Blood.* vol. 98, no 6, p. 1949-1954.

Kawada, H., T. Ito, P. N. Pharr, D. D. Spyropoulos, D. K. Watson et M. Ogawa. 2001. «Defective megakaryopoiesis and abnormal erythroid development in Fli-1 gene-targeted mice». *Int J Hematol.* vol. 73, no 4, p. 463-468.

- Khimani, A. H., M. Lim, T. G. Graf, T. F. Smith et R. M. Ruprecht. 1997. «Phylogenetic relationship of the complete Rauscher murine leukemia virus genome with other murine leukemia virus genomes». *Virology*. vol. 238, no 1, p. 64-67.
- Kiel, M. J., O. H. Yilmaz, T. Iwashita, O. H. Yilmaz, C. Terhorst et S. J. Morrison. 2005. «SLAM family receptors distinguish hematopoietic stem and progenitor cells and reveal endothelial niches for stem cells». *Cell*. vol. 121, no 7, p. 1109-1121.
- Kihm, A. J., Y. Kong, W. Hong, J. E. Russell, S. Rouda, K. Adachi, M. C. Simon, G. A. Blobel et M. J. Weiss. 2002. «An abundant erythroid protein that stabilizes free alpha-haemoglobin». *Nature*. vol. 417, no 6890, p. 758-763.
- Kilbey, A., H. Alzuherri, J. McColl, C. Cales, J. Frampton et C. Bartholomew. 2005. «The Evl1 proto-oncoprotein blocks endomitosis in megakaryocytes by inhibiting sustained cyclin-dependent kinase 2 catalytic activity». *Br J Haematol*. vol. 130, no 6, p. 902-911.
- Kim, I., S. He, O. H. Yilmaz, M. J. Kiel et S. J. Morrison. 2006. «Enhanced purification of fetal liver hematopoietic stem cells using SLAM family receptors». *Blood*. vol. 108, no 2, p. 737-744.
- Kim, J. A., Y. J. Jung, J. Y. Seoh, S. Y. Woo, J. S. Seo et H. L. Kim. 2002. «Gene expression profile of megakaryocytes from human cord blood CD34(+) cells ex vivo expanded by thrombopoietin». *Stem Cells*. vol. 20, no 5, p. 402-416.
- Kina, T., K. Ikuta, E. Takayama, K. Wada, A. S. Majumdar, I. L. Weissman et Y. Katsura. 2000. «The monoclonal antibody TER-119 recognizes a molecule associated with glycophorin A and specifically marks the late stages of murine erythroid lineage». *Br J Haematol*. vol. 109, no 2, p. 280-287.
- Kirchgessner, H., J. Dietrich, J. Scherer, P. Isomaki, V. Korinek, I. Hilgert, E. Bruyns, A. Leo, A. P. Cope et B. Schraven. 2001. «The transmembrane adaptor protein TRIM regulates T cell receptor (TCR) expression and TCR-mediated signaling via an association with the TCR zeta chain». *J Exp Med*. vol. 193, no 11, p. 1269-1284.
- Kiss, R. S., Z. Ma, K. Nakada-Tsukui, E. Brugnera, G. Vassiliou, H. M. McBride, K. S. Ravichandran et Y. L. Marcel. 2006. «The lipoprotein receptor-related protein-1 (LRP) adapter protein GULP mediates trafficking of the LRP ligand prosaposin, leading to sphingolipid and free cholesterol accumulation in late endosomes and impaired efflux». *J Biol Chem*. vol. 281, no 17, p. 12081-12092.
- Kissel, K., S. Berber, A. Nockher, S. Santoso, G. Bein et H. Hackstein. 2006. «Human platelets target dendritic cell differentiation and production of proinflammatory cytokines». *Transfusion*. vol. 46, no 5, p. 818-827.
- Kitajima, K., M. Tanaka, J. Zheng, H. Yen, A. Sato, D. Sugiyama, H. Umehara, E. Sakai et T. Nakano. 2006. «Redirecting differentiation of hematopoietic progenitors by a transcription factor, GATA-2». *Blood*. vol. 107, no 5, p. 1857-1863.

- Kitamura, T., T. Tange, T. Terasawa, S. Chiba, T. Kuwaki, K. Miyagawa, Y. F. Piao, K. Miyazono, A. Urabe et F. Takaku. 1989. «Establishment and characterization of a unique human cell line that proliferates dependently on GM-CSF, IL-3, or erythropoietin». *J Cell Physiol.* vol. 140, no 2, p. 323-334.
- Kluger, Y., H. Yu, J. Qian et M. Gerstein. 2003. «Relationship between gene co-expression and probe localization on microarray slides». *BMC Genomics.* vol. 4, no 1, p. 49.
- Kogan, S. C., J. M. Ward, M. R. Anver, J. J. Berman, C. Brayton, R. D. Cardiff, J. S. Carter, S. de Coronado, J. R. Downing, T. N. Fredrickson, D. C. Haines, A. W. Harris, N. L. Harris, H. Hiai, E. S. Jaffe, I. C. MacLennan, P. P. Pandolfi, P. K. Pattengale, A. S. Perkins, R. M. Simpson, M. S. Tuttle, J. F. Wong et H. C. Morse, 3rd. 2002. «Bethesda proposals for classification of nonlymphoid hematopoietic neoplasms in mice». *Blood.* vol. 100, no 1, p. 238-245.
- Kohl, R., M. Antoine, K. Reimers et P. Kiefer. 2002. «FGF3 attached to a phospholipid membrane anchor gains a high transforming capacity. Implications of microdomains for FGF3 cell transformation». *J Biol Chem.* vol. 277, no 36, p. 32760-32767.
- Kohonen, T. 1999. «Comparison of SOM point densities based on different criteria». *Neural Comput.* vol. 11, no 8, p. 2081-2095.
- Korutla, L., P. J. Wang, D. M. Lewis, J. H. Neustadter, M. F. Stromberg et S. A. Mackler. 2002. «Differences in expression, actions and cocaine regulation of two isoforms for the brain transcriptional regulator NAC1». *Neuroscience.* vol. 110, no 3, p. 421-429.
- Korz, C., A. Pscherer, A. Benner, D. Mertens, C. Schaffner, E. Leupolt, H. Dohner, S. Stilgenbauer et P. Lichter. 2002. «Evidence for distinct pathomechanisms in B-cell chronic lymphocytic leukemia and mantle cell lymphoma by quantitative expression analysis of cell cycle and apoptosis-associated genes». *Blood.* vol. 99, no 12, p. 4554-4561.
- Kozak, C. A., et A. Chakraborti. 1996. «Single amino acid changes in the murine leukemia virus capsid protein gene define the target of Fv1 resistance». *Virology.* vol. 225, no 2, p. 300-305.
- Kozak, C. A., et R. R. O'Neill. 1987. «Diverse wild mouse origins of xenotropic, mink cell focus-forming, and two types of ecotropic proviral genes». *J Virol.* vol. 61, no 10, p. 3082-3088.
- Kransdorf, E. P., S. Z. Wang, S. Z. Zhu, T. B. Langston, J. W. Rupon et G. D. Ginder. 2006. «MBD2 is a critical component of a methyl cytosine-binding protein complex isolated from primary erythroid cells». *Blood.* vol. 108, no 8, p. 2836-2845.
- Krishnamurthy, P. C., G. Du, Y. Fukuda, D. Sun, J. Sampath, K. E. Mercer, J. Wang, B. Sosa-Pineda, K. G. Murti et J. D. Schuetz. 2006. «Identification of a mammalian mitochondrial porphyrin transporter». *Nature.* vol. 443, no 7111, p. 586-589.

- Kubota, H., H. Okazaki, M. Onuma, S. Kano, M. Hattori et N. Minato. 1990. «Identification and gene cloning of a new phosphatidylinositol-linked antigen expressed on mature lymphocytes. Down-regulation by lymphocyte activation». *J Immunol.* vol. 145, no 11, p. 3924-3931.
- Kuhner, M. K., et J. Felsenstein. 1994. «A simulation comparison of phylogeny algorithms under equal and unequal evolutionary rates». *Mol Biol Evol.* vol. 11, no 3, p. 459-468.
- Lagadec, P., O. Dejoux, M. Ticchioni, F. Cottrez, M. Johansen, E. J. Brown et A. Bernard. 2003. «Involvement of a CD47-dependent pathway in platelet adhesion on inflamed vascular endothelium under flow». *Blood.* vol. 101, no 12, p. 4836-4843.
- Lagasse, E., et I. L. Weissman. 1996. «Flow cytometric identification of murine neutrophils and monocytes». *J Immunol Methods.* vol. 197, no 1-2, p. 139-150.
- Lahlil, R., E. Lecuyer, S. Herblot et T. Hoang. 2004. «SCL assembles a multifactorial complex that determines glycophorin A expression». *Mol Cell Biol.* vol. 24, no 4, p. 1439-1452..
- Lanza, F. 2006. «Bernard-Soulier syndrome (hemorrhagicparous thrombocytic dystrophy)». *Orphanet J Rare Dis.* vol. 1, p. 46.
- Lassaux, A., M. Sitbon et J. L. Battini. 2005. «Residues in the murine leukemia virus capsid that differentially govern resistance to mouse Fv1 and human Ref1 restrictions». *J Virol.* vol. 79, no 10, p. 6560-6564.
- Lau, L. M., J. L. Wee, M. D. Wright, G. W. Moseley, P. M. Hogarth, L. K. Ashman et D. E. Jackson. 2004. «The tetraspanin superfamily member CD151 regulates outside-in integrin α IIb β 3 signaling and platelet function». *Blood.* vol. 104, no 8, p. 2368-2375.
- LeBien, T. W. 2000. «Fates of human B-cell precursors». *Blood.* vol. 96, no 1, p. 9-23.
- Lecuyer, E., et T. Hoang. 2004. «SCL: from the origin of hematopoiesis to stem cells and leukemia». *Exp Hematol.* vol. 32, no 1, p. 11-24.
- Ledbetter, J. A., R. V. Rouse, H. S. Micklem et L. A. Herzenberg. 1980. «T cell subsets defined by expression of Lyt-1,2,3 and Thy-1 antigens. Two-parameter immunofluorescence and cytotoxicity analysis with monoclonal antibodies modifies current views». *J Exp Med.* vol. 152, no 2, p. 280-295.
- Lengyel, E., K. Sawada et R. Salgia. 2007. «Tyrosine kinase mutations in human cancer». *Curr Mol Med.* vol. 7, no 1, p. 77-84.
- Lennon, P. A., L. V. Abruzzo, L. J. Medeiros, C. Cromwell, X. Zhang, C. C. Yin, S. M. Kornblau, M. Konopieva et P. Lin. 2007. «Aberrant EVII expression in acute myeloid leukemias associated with the t(3;8)(q26;q24)». *Cancer Genet Cytogenet.* vol. 177, no 1, p. 37-42.

- Lewinski, M. K., M. Yamashita, M. Emerman, A. Ciuffi, H. Marshall, G. Crawford, F. Collins, P. Shinn, J. Leipzig, S. Hannehalli, C. C. Berry, J. R. Ecker et F. D. Bushman. 2006. «Retroviral DNA integration: viral and cellular determinants of target-site selection». *PLoS Pathog.* vol. 2, no 6, p. e60.
- Lewis, J. D., L. A. Payton, J. G. Whitford, J. A. Byrne, D. I. Smith, L. Yang et R. K. Bright. 2007. «Induction of tumorigenesis and metastasis by the murine orthologue of tumor protein D52». *Mol Cancer Res.* vol. 5, no 2, p. 133-144.
- Li, C., et W. H. Wong. 2001. «Model-based analysis of oligonucleotide arrays: expression index computation and outlier detection». *Proc Natl Acad Sci U S A.* vol. 98, no 1, p. 31-36..
- Li, J. J., Y. Q. Huang, R. Basch et S. Karparkin. 2001. «Thrombin induces the release of angiotensin-1 from platelets». *Thromb Haemost.* vol. 85, no 2, p. 204-206..
- Li, Y., E. Golemis, J. W. Hartley et N. Hopkins. 1987. «Disease specificity of nondefective Friend and Moloney murine leukemia viruses is controlled by a small number of nucleotides». *J Virol.* vol. 61, no 3, p. 693-700.
- Lim, K. H., A. T. Baines, J. J. Fiordalisi, M. Shipitsin, L. A. Feig, A. D. Cox, C. J. Der et C. M. Counter. 2005. «Activation of RalA is critical for Ras-induced tumorigenesis of human cells». *Cancer Cell.* vol. 7, no 6, p. 533-545.
- Liu, Y., D. Promeneur, A. Rojek, N. Kumar, J. Frokiaer, S. Nielsen, L. S. King, P. Agre et J. M. Carbrey. 2007. «Aquaporin 9 is the major pathway for glycerol uptake by mouse erythrocytes, with implications for malarial virulence». *Proc Natl Acad Sci U S A.* vol. 104, no 30, p. 12560-12564.
- Loffredo, G., L. Baronciani, P. Noris, F. Menna, A. B. Federici et C. L. Balduini. 2006. «von Willebrand disease type 2B must be always considered in the differential diagnosis of genetic thrombocytopenias with giant platelets». *Platelets.* vol. 17, no 3, p. 149-152.
- Lok, C. N., et T. T. Loh. 1998. «Regulation of transferrin function and expression: review and update». *Biol Signals Recept.* vol. 7, no 3, p. 157-178.
- Long, M. W., et N. Williams. 1982. «Differences in the regulation of megakaryocytopoiesis in the murine bone marrow and spleen». *Leuk Res.* vol. 6, no 5, p. 721-728..
- Longhurst, C. M., M. M. White, D. A. Wilkinson et L. K. Jennings. 1999. «A CD9, alphaIIb beta3, integrin-associated protein, and GPIb/V/IX complex on the surface of human platelets is influenced by alphaIIb beta3 conformational states». *Eur J Biochem.* vol. 263, no 1, p. 104-111.
- Loose, M., G. Swiers et R. Patient. 2007. «Transcriptional networks regulating hematopoietic cell fate decisions». *Curr Opin Hematol.* vol. 14, no 4, p. 307-314.

- Lu, M., H. Kawamoto, Y. Katsube, T. Ikawa et Y. Katsura. 2002. «The common myelolymphoid progenitor: a key intermediate stage in hemopoiesis generating T and B cells». *J Immunol.* vol. 169, no 7, p. 3519-3525.
- Lund, A. H., et F. S. Pedersen. 1999. «The nucleotide sequence of the high-leukemogenic murine retrovirus SL3-3 reveals a patch of mink cell focus forming-like sequences upstream of the ecotropic envelope gene. Brief report». *Arch Virol.* vol. 144, no 11, p. 2207-2212.
- Ma, S. L., A. B. Sorensen, S. Kunder, K. D. Sorensen, L. Quintanilla-Martinez, D. W. Morris, J. Schmidt et F. S. Pedersen. 2006. «The Icsbp locus is a common proviral insertion site in mature B-cell lymphomas/plasmacytomas induced by exogenous murine leukemia virus». *Virology.* vol. 352, no 2, p. 306-318.
- Maki, K., T. Yamagata, I. Yamazaki, H. Oda et K. Mitani. 2006. «Development of megakaryoblastic leukaemia in Runx1-Evi1 knock-in chimaeric mouse». *Leukemia.* vol. 20, no 8, p. 1458-1460.
- Malin, D., I. M. Kim, E. Boettcher, T. V. Kalin, S. Ramakrishna, L. Meliton, V. Ustiyan, X. Zhu et V. V. Kalinichenko. 2007. «Forkhead box F1 is essential for migration of mesenchymal cells and directly induces integrin-beta3 expression». *Mol Cell Biol.* vol. 27, no 7, p. 2486-2498.
- Manley, N. R., M. A. O'Connell, P. A. Sharp et N. Hopkins. 1989. «Nuclear factors that bind to the enhancer region of nondefective Friend murine leukemia virus». *J Virol.* vol. 63, no 10, p. 4210-4223.
- Manley, N. R., M. O'Connell, W. Sun, N. A. Speck et N. Hopkins. 1993. «Two factors that bind to highly conserved sequences in mammalian type C retroviral enhancers». *J Virol.* vol. 67, no 4, p. 1967-1975.
- Marini, A., A. Mirmohammadsadegh, S. Nambiar, A. Gustrau, T. Ruzicka et U. R. Hengge. 2006. «Epigenetic inactivation of tumor suppressor genes in serum of patients with cutaneous melanoma». *J Invest Dermatol.* vol. 126, no 2, p. 422-431.
- Marjanovic, J. A., Z. Li, A. Stojanovic et X. Du. 2005. «Stimulatory roles of nitric-oxide synthase 3 and guanylyl cyclase in platelet activation». *J Biol Chem.* vol. 280, no 45, p. 37430-37438.
- Martel, V., C. Racaud-Sultan, S. Dupe, C. Marie, F. Paulhe, A. Galmiche, M. R. Block et C. Albiges-Rizo. 2001. «Conformation, localization, and integrin binding of talin depend on its interaction with phosphoinositides». *J Biol Chem.* vol. 276, no 24, p. 21217-21227.
- Martin, D. I., L. I. Zon, G. Mutter et S. H. Orkin. 1990. «Expression of an erythroid transcription factor in megakaryocytic and mast cell lineages». *Nature.* vol. 344, no 6265, p. 444-447.
- Marty, C., D. D. Browning et R. D. Ye. 2003. «Identification of tetratricopeptide repeat 1 as an adaptor protein that interacts with heterotrimeric G proteins and the small GTPase Ras». *Mol Cell Biol.* vol. 23, no 11, p. 3847-3858.

- Martz, E. 2002. «Protein Explorer: easy yet powerful macromolecular visualization». *Trends Biochem Sci.* vol. 27, no 2, p. 107-109.
- Matsumoto-Taniura, N., F. Pirollet, R. Monroe, L. Gerace et J. M. Westendorf. 1996. «Identification of novel M phase phosphoproteins by expression cloning». *Mol Biol Cell.* vol. 7, no 9, p. 1455-1469.
- Matsunaga, T., I. Tanaka, M. Kobune, Y. Kawano, M. Tanaka, K. Kuribayashi, S. Iyama, T. Sato, Y. Sato, R. Takimoto, T. Takayama, J. Kato, T. Ninomiya, H. Hamada et Y. Niitsu. 2006. «Ex vivo large-scale generation of human platelets from cord blood CD34+ cells». *Stem Cells.* vol. 24, no 12, p. 2877-2887.
- Mattera, R., Y. C. Tsai, A. M. Weissman et J. S. Bonifacino. 2006. «The Rab5 guanine nucleotide exchange factor Rabex-5 binds ubiquitin (Ub) and functions as a Ub ligase through an atypical Ub-interacting motif and a zinc finger domain». *J Biol Chem.* vol. 281, no 10, p. 6874-6883.
- McClung, J. K., E. R. Jupe, X. T. Liu et R. T. Dell'Orco. 1995. «Prohibitin: potential role in senescence, development, and tumor suppression». *Exp Gerontol.* vol. 30, no 2, p. 99-124.
- McDonald, T. P., et P. S. Sullivan. 1993. «Megakaryocytic and erythrocytic cell lines share a common precursor cell». *Exp Hematol.* vol. 21, no 10, p. 1316-1320.
- McNaghy, K. M., et T. Graf. 2003. «E26 leukemia virus converts primitive erythroid cells into cycling multilineage progenitors». *Blood.* vol. 101, no 3, p. 1103-1110.
- Melikyan, G. B., R. M. Markosyan, S. A. Brener, Y. Rozenberg et F. S. Cohen. 2000. «Role of the cytoplasmic tail of ecotropic moloney murine leukemia virus Env protein in fusion pore formation». *J Virol.* vol. 74, no 1, p. 447-455.
- Meri, S., B. P. Morgan, A. Davies, R. H. Daniels, M. G. Olavesen, H. Waldmann et P. J. Lachmann. 1990. «Human protectin (CD59), an 18,000-20,000 MW complement lysis restricting factor, inhibits C5b-8 catalysed insertion of C9 into lipid bilayers». *Immunology.* vol. 71, no 1, p. 1-9.
- Merregaert, J., M. Janowski et E. P. Reddy. 1987. «Nucleotide sequence of a radiation leukemia virus genome». *Virology.* vol. 158, no 1, p. 88-102.
- Mikkers, H., J. Allen, P. Knipscheer, L. Romeijn, A. Hart, E. Vink et A. Berns. 2002. «High-throughput retroviral tagging to identify components of specific signaling pathways in cancer». *Nat Genet.* vol. 32, no 1, p. 153-159.
- Misra, J., W. Schmitt, D. Hwang, L. L. Hsiao, S. Gullans, G. Stephanopoulos et G. Stephanopoulos. 2002. «Interactive exploration of microarray gene expression patterns in a reduced dimensional space». *Genome Res.* vol. 12, no 7, p. 1112-1120.

- Miyashita, H., M. Kanemura, T. Yamazaki, M. Abe et Y. Sato. 2004. «Vascular endothelial zinc finger 1 is involved in the regulation of angiogenesis: possible contribution of stathmin/OP18 as a downstream target gene». *Arterioscler Thromb Vasc Biol.* vol. 24, no 5, p. 878-884.
- Moloney, J. B. 1960. «Biological studies on a lymphoid-leukemia virus extracted from sarcoma 37. I. Origin and introductory investigations». *J Natl Cancer Inst.* vol. 24, p. 933-951.
- Moon, Y. A., N. A. Shah, S. Mohapatra, J. A. Warrington et J. D. Horton. 2001. «Identification of a mammalian long chain fatty acyl elongase regulated by sterol regulatory element-binding proteins». *J Biol Chem.* vol. 276, no 48, p. 45358-45366.
- Morerio, C., M. Acquila, C. Rosanda, A. Rapella, C. Dufour, F. Locatelli, E. Maserati, F. Pasquali et C. Panarello. 2004. «HCMOGT-1 is a novel fusion partner to PDGFRB in juvenile myelomonocytic leukemia with t(5;17)(q33;p11.2)». *Cancer Res.* vol. 64, no 8, p. 2649-2651.
- Morse, H. C., 3rd, M. R. Anver, T. N. Fredrickson, D. C. Haines, A. W. Harris, N. L. Harris, E. S. Jaffe, S. C. Kogan, I. C. MacLennan, P. K. Pattengale et J. M. Ward. 2002. «Bethesda proposals for classification of lymphoid neoplasms in mice». *Blood.* vol. 100, no 1, p. 246-258.
- Moskow, J. J., F. Bullrich, K. Huebner, I. O. Daar, A. M. Buchberg, A. Mamo, J. Kros, E. Kroon, J. Bijl, A. Thompson, N. Mayotte, S. Girard, R. Bisailon, N. Beslu, M. Featherstone, G. Sauvageau, C. P. Chang, Y. Jacobs, T. Nakamura, N. A. Jenkins, N. G. Copeland, M. L. Cleary, V. Azcoitia, M. Aracil, A. C. Martinez et M. Torres. 1995. «Meis1, a PBX1-related homeobox gene involved in myeloid leukemia in BXH-2 mice Molecular dissection of Meis1 reveals 2 domains required for leukemia induction and a key role for Hoxa gene activation
Meis proteins are major in vivo DNA binding partners for wild-type but not chimeric Pbx proteins
The homeodomain protein Meis1 is essential for definitive hematopoiesis and vascular patterning in the mouse embryo». *Mol Cell Biol.* vol. 15, no 10, p. 5434-5443.
- Mossuz, P., M. H. Prandini, D. Leroux, O. Valiron, A. Schweitzer, B. Sasseti et L. Kolodie. 1997. «Effects of retinoic acid on a new human erythromegakaryocytic cell line AP-217». *Leuk Res.* vol. 21, no 6, p. 529-537.
- Nakai, M., M. Fujita, M. Waragai, S. Sugama, J. Wei, H. Akatsu, C. Ohtaka-Maruyama, H. Okado et M. Hashimoto. 2007. «Expression of alpha-synuclein, a presynaptic protein implicated in Parkinson's disease, in erythropoietic lineage». *Biochem Biophys Res Commun.* vol. 358, no 1, p. 104-110.
- Nakajima, H., et M. Colonna. 2000. «2B4: an NK cell activating receptor with unique specificity and signal transduction mechanism». *Hum Immunol.* vol. 61, no 1, p. 39-43.
- Nakayama, K., N. Nakayama, B. Davidson, J. J. Sheu, N. Jinawath, A. Santillan, R. Salani, R. E. Bristow, P. J. Morin, R. J. Kurman, T. L. Wang et M. Shih Ie. 2006. «A

- BTB/POZ protein, NAC-1, is related to tumor recurrence and is essential for tumor growth and survival». *Proc Natl Acad Sci U S A.* vol. 103, no 49, p. 18739-18744.
- Nakorn, T. N., T. Miyamoto et I. L. Weissman. 2003. «Characterization of mouse clonogenic megakaryocyte progenitors». *Proc Natl Acad Sci U S A.* vol. 100, no 1, p. 205-210.
- Nasdala, I., K. Wolburg-Buchholz, H. Wolburg, A. Kuhn, K. Ebnet, G. Brachtendorf, U. Samulowitz, B. Kuster, B. Engelhardt, D. Vestweber et S. Butz. 2002. «A transmembrane tight junction protein selectively expressed on endothelial cells and platelets». *J Biol Chem.* vol. 277, no 18, p. 16294-16303.
- Neame, P. B., P. Soamboonsrup, G. Browman, R. D. Barr, N. Saeed, B. Chan, M. Pai, A. Bengier, W. E. Wilson, I. R. Walker et al. 1985. «Simultaneous or sequential expression of lymphoid and myeloid phenotypes in acute leukemia». *Blood.* vol. 65, no 1, p. 142-148.
- Nilsson, N. E., K. Kotarsky, C. Owman et B. Olde. 2003. «Identification of a free fatty acid receptor, FFA2R, expressed on leukocytes and activated by short-chain fatty acids». *Biochem Biophys Res Commun.* vol. 303, no 4, p. 1047-1052.
- Nouvel, P., H. Philippe, H. Condamine et J. J. Panthier. 1993. «Analysis of variability among endogenous ecotropic MuLV loci in laboratory mice». *Virology.* vol. 193, no 1, p. 450-455.
- Nucifora, G., L. Laricchia-Robbio et V. Senyuk. 2006. «EVI1 and hematopoietic disorders: history and perspectives». *Gene.* vol. 368, p. 1-11.
- O'Neill, R. R., C. E. Buckler, T. S. Theodore, M. A. Martin et R. Repaske. 1985. «Envelope and long terminal repeat sequences of a cloned infectious NZB xenotropic murine leukemia virus». *J Virol.* vol. 53, no 1, p. 100-106.
- Okada, Y., R. Nagai, T. Sato, E. Matsuura, T. Minami, I. Morita et T. Doi. 2003. «Homeodomain proteins MEIS1 and PBXs regulate the lineage-specific transcription of the platelet factor 4 gene». *Blood.* vol. 101, no 12, p. 4748-4756.
- Okimoto, M. A., et H. Fan. 1999. «Identification of directly infected cells in the bone marrow of neonatal moloney murine leukemia virus-infected mice by use of a moloney murine leukemia virus-based vector». *J Virol.* vol. 73, no 2, p. 1617-1623.
- Oklu, R., et R. Hesketh. 2000. «The latent transforming growth factor beta binding protein (LTBP) family». *Biochem J.* vol. 352 Pt 3, p. 601-610.
- Oliver, A. M., F. Martin et J. F. Kearney. 1997. «Mouse CD38 is down-regulated on germinal center B cells and mature plasma cells». *J Immunol.* vol. 158, no 3, p. 1108-1115.
- Olweus, J., F. Lund-Johansen et V. Horejsi. 1993. «CD53, a protein with four membrane-spanning domains, mediates signal transduction in human monocytes and B cells». *J Immunol.* vol. 151, no 2, p. 707-716.

- Oosterwegel, M., M. van de Wetering, D. Dooijes, L. Klomp, A. Winoto, K. Georgopoulos, F. Meijlink et H. Clevers. 1991. «Cloning of murine TCF-1, a T cell-specific transcription factor interacting with functional motifs in the CD3-epsilon and T cell receptor alpha enhancers». *J Exp Med.* vol. 173, no 5, p. 1133-1142.
- Orkin, S. H. 1992. «GATA-binding transcription factors in hematopoietic cells». *Blood.* vol. 80, no 3, p. 575-581.
- Ott, D. E., J. Keller, K. Sill et A. Rein. 1992. «Phenotypes of murine leukemia virus-induced tumors: influence of 3' viral coding sequences». *J Virol.* vol. 66, no 10, p. 6107-6116.
- Ott, D., R. Friedrich et A. Rein. 1990. «Sequence analysis of amphotropic and 10A1 murine leukemia viruses: close relationship to mink cell focus-inducing viruses». *J Virol.* vol. 64, no 2, p. 757-766.
- Ouyang, W., N. G. Jacobson, D. Bhattacharya, J. D. Gorham, D. Fenoglio, W. C. Sha, T. L. Murphy et K. M. Murphy. 1999. «The Ets transcription factor ERM is Th1-specific and induced by IL-12 through a Stat4-dependent pathway». *Proc Natl Acad Sci U S A.* vol. 96, no 7, p. 3888-3893.
- Overholtzer, M., J. Zhang, G. A. Smolen, B. Muir, W. Li, D. C. Sgroi, C. X. Deng, J. S. Brugge et D. A. Haber. 2006. «Transforming properties of YAP, a candidate oncogene on the chromosome 11q22 amplicon». *Proc Natl Acad Sci U S A.* vol. 103, no 33, p. 12405-12410.
- Pagano, L., A. Pulsoni, M. Vignetti, L. Mele, L. Fianchi, M. C. Petti, S. Mirto, P. Falucci, P. Fazi, G. Broccia, G. Specchia, F. Di Raimondo, L. Pacilli, P. Leoni, S. Ladogana, E. Gallo, A. Venditti, G. Avanzi, A. Camera, V. Liso, G. Leone et F. Mandelli. 2002. «Acute megakaryoblastic leukemia: experience of GIMEMA trials». *Leukemia.* vol. 16, no 9, p. 1622-1626.
- Palis, J., S. Robertson, M. Kennedy, C. Wall et G. Keller. 1999. «Development of erythroid and myeloid progenitors in the yolk sac and embryo proper of the mouse». *Development.* vol. 126, no 22, p. 5073-5084.
- Pan, H. W., H. Y. Chou, S. H. Liu, S. Y. Peng, C. L. Liu et H. C. Hsu. 2006. «Role of L2DTL, cell cycle-regulated nuclear and centrosome protein, in aggressive hepatocellular carcinoma». *Cell Cycle.* vol. 5, no 22, p. 2676-2687.
- Panet, A., et H. Cedar. 1977. «Selective degradation of integrated murine leukemia proviral DNA by deoxyribonucleases». *Cell.* vol. 11, no 4, p. 933-940.
- Paoletti, F., A. M. Vannucchi, A. Mocali, R. Caporale et S. A. Burstein. 1995. «Identification and conditions for selective expression of megakaryocytic markers in Friend erythroleukemia cells». *Blood.* vol. 86, no 7, p. 2624-2631.
- Papayannopoulou, T., M. Brice, D. Farrer et K. Kaushansky. 1996. «Insights into the cellular mechanisms of erythropoietin-thrombopoietin synergy». *Exp Hematol.* vol. 24, no 5, p. 660-669.

- Passegue, E., et I. L. Weisman. 2005. «Leukemic stem cells: where do they come from?». *Stem Cell Rev.* vol. 1, no 3, p. 181-188.
- Patel, S. R., J. H. Hartwig et J. E. Italiano, Jr. 2005. «The biogenesis of platelets from megakaryocyte proplatelets». *J Clin Invest.* vol. 115, no 12, p. 3348-3354
- Pedersen, F. S., R. L. Crowther, D. Y. Tenney, A. M. Reimold et W. A. Haseltine. 1981. «Novel leukaemogenic retroviruses isolated from cell line derived from spontaneous AKR tumour». *Nature.* vol. 292, no 5819, p. 167-170.
- Perbal, B., C. Martinerie, R. Sainson, M. Werner, B. He et B. Roizman. 1999. «The C-terminal domain of the regulatory protein NOVH is sufficient to promote interaction with fibulin 1C: a clue for a role of NOVH in cell-adhesion signaling». *Proc Natl Acad Sci U S A.* vol. 96, no 3, p. 869-874.
- Pereira, R., C. T. Quang, I. Lesault, H. Dolznig, H. Beug et J. Ghysdael. 1999. «FLI-1 inhibits differentiation and induces proliferation of primary erythroblasts». *Oncogene.* vol. 18, no 8, p. 1597-1608.
- Perentesis, J., N. K. Ramsay, R. Brunning, J. H. Kersey et A. H. Filipovich. 1983. «Biphenotypic leukemia: immunologic and morphologic evidence for a common lymphoid-myeloid progenitor in humans». *J Pediatr.* vol. 102, no 1, p. 63-67.
- Perry, C., et H. Soreq. 2002. «Transcriptional regulation of erythropoiesis. Fine tuning of combinatorial multi-domain elements». *Eur J Biochem.* vol. 269, no 15, p. 3607-3618.
- Perryman, S., J. Nishio et B. Chesebro. 1991. «Complete nucleotide sequence of Friend murine leukemia virus, strain FB29». *Nucleic Acids Res.* vol. 19, no 24, p. 6950.
- Peters, L. L., R. A. Swearingen, S. G. Andersen, B. Gwynn, A. J. Lambert, R. Li, S. E. Lux et G. A. Churchill. 2004. «Identification of quantitative trait loci that modify the severity of hereditary spherocytosis in wan, a new mouse model of band-3 deficiency». *Blood.* vol. 103, no 8, p. 3233-3240.
- Pevny, L., M. C. Simon, E. Robertson, W. H. Klein, S. F. Tsai, V. D'Agati, S. H. Orkin et F. Costantini. 1991. «Erythroid differentiation in chimaeric mice blocked by a targeted mutation in the gene for transcription factor GATA-1». *Nature.* vol. 349, no 6306, p. 257-260.
- Pflugh, David L., Stephen E. Maher et Alfred L. M. Bothwell (2002). *Ly-6 Superfamily Members Ly-6A/E, Ly-6C, and Ly-6I Recognize Two Potential Ligands Expressed by B Lymphocytes.* 169: 5130-5136 p
- Pinto, A., L. Del Vecchio, A. Carbone, M. Roncadin, R. Volpe, D. Serraino, S. Monfardini, A. Colombatti et V. Zagonel. 1991. «Expression of myelomonocytic antigens is associated with unfavourable clinicoprognostic factors in B-cell chronic lymphocytic leukaemia». *Ann Oncol.* vol. 2 Suppl 2, p. 107-113.

- Pohl, U., J. S. Smith, I. Tachibana, K. Ueki, H. K. Lee, S. Ramaswamy, Q. Wu, H. W. Mohrenweiser, R. B. Jenkins et D. N. Louis. 2000. «EHD2, EHD3, and EHD4 encode novel members of a highly conserved family of EH domain-containing proteins». *Genomics*. vol. 63, no 2, p. 255-262.
- Poiesz, B. J., F. W. Ruscetti, A. F. Gazdar, P. A. Bunn, J. D. Minna et R. C. Gallo. 1980. «Detection and isolation of type C retrovirus particles from fresh and cultured lymphocytes of a patient with cutaneous T-cell lymphoma». *Proc Natl Acad Sci U S A*. vol. 77, no 12, p. 7415-7419.
- Poliquin, L., D. Bergeron, J. L. Fortier, Y. Paquette, R. Bergeron et E. Rassart. 1992. «Determinants of thymotropism in Kaplan radiation leukemia virus and nucleotide sequence of its envelope region». *J Virol*. vol. 66, no 8, p. 5141-5146.
- Poulos, T. L. 2006. «Soluble guanylate cyclase». *Curr Opin Struct Biol*. vol. 16, no 6, p. 736-743.
- Pupa, S. M., S. Giuffre, F. Castiglioni, L. Bertola, M. Cantu, I. Bongarzone, P. Baldassari, R. Mortarini, W. S. Argraves, A. Anichini, S. Menard et E. Tagliabue. 2007. «Regulation of breast cancer response to chemotherapy by fibulin-1». *Cancer Res*. vol. 67, no 9, p. 4271-4277.
- Qing, J., V. M. Maher, H. Tran, W. S. Argraves, R. W. Dunstan et J. J. McCormick. 1997. «Suppression of anchorage-independent growth and matrigel invasion and delayed tumor formation by elevated expression of fibulin-1D in human fibrosarcoma-derived cell lines». *Oncogene*. vol. 15, no 18, p. 2159-2168.
- Quesenberry, P. J. 2006. «The continuum model of marrow stem cell regulation». *Curr Opin Hematol*. vol. 13, no 4, p. 216-221.
- Raisch, K. P., M. Pizzato, H. Y. Sun, Y. Takeuchi, L. W. Cashdollar et S. E. Grossberg. 2003. «Molecular cloning, complete sequence, and biological characterization of a xenotropic murine leukemia virus constitutively released from the human B-lymphoblastoid cell line DG-75». *Virology*. vol. 308, no 1, p. 83-91.
- Rajpal, A., Y. A. Cho, B. Yelent, P. H. Koza-Taylor, D. Li, E. Chen, M. Whang, C. Kang, T. G. Turi et A. Winoto. 2003. «Transcriptional activation of known and novel apoptotic pathways by Nur77 orphan steroid receptor». *Embo J*. vol. 22, no 24, p. 6526-6536.
- Raman, C. 2002. «CD5, an important regulator of lymphocyte selection and immune tolerance». *Immunol Res*. vol. 26, no 1-3, p. 255-263.
- Rasheed, S., B. K. Pal et M. B. Gardner. 1982. «Characterization of a highly oncogenic murine leukemia virus from wild mice». *Int J Cancer*. vol. 29, no 3, p. 345-350.
- Rassart, E., M. Shang, Y. Boie et P. Jolicoeur. 1986. «Studies on emerging radiation leukemia virus variants in C57BL/Ka mice». *J Virol*. vol. 58, no 1, p. 96-106.
- Rauscher, F. J. 1962. «A virus-induced disease of mice characterized by erythrocytopoiesis and lymphoid leukemia». *J Natl Cancer Inst*. vol. 29, p. 515-543.

- Rein, A., C. Yang, J. A. Haynes, J. Mirro et R. W. Compans. 1998. «Evidence for cooperation between murine leukemia virus Env molecules in mixed oligomers». *J Virol.* vol. 72, no 4, p. 3432-3435.
- Richmond, T. D., M. Chohan et D. L. Barber. 2005. «Turning cells red: signal transduction mediated by erythropoietin». *Trends Cell Biol.* vol. 15, no 3, p. 146-155.
- Riddle, E. L., R. A. Schwartzman, M. Bond et P. A. Insel. 2005. «Multi-tasking RGS proteins in the heart: the next therapeutic target?». *Circ Res.* vol. 96, no 4, p. 401-411.
- Roark, E. F., D. R. Keene, C. C. Haudenschild, S. Godyna, C. D. Little et W. S. Argraves. 1995. «The association of human fibulin-1 with elastic fibers: an immunohistological, ultrastructural, and RNA study». *J Histochem Cytochem.* vol. 43, no 4, p. 401-411.
- Robert-Moreno, A., L. Espinosa, J. L. de la Pompa et A. Bigas. 2005. «RBPjkappa-dependent Notch function regulates Gata2 and is essential for the formation of intra-embryonic hematopoietic cells». *Development.* vol. 132, no 5, p. 1117-1126.
- Romero, X., D. Benitez, S. March, R. Vilella, M. Miralpeix et P. Engel. 2004. «Differential expression of SAP and EAT-2-binding leukocyte cell-surface molecules CD84, CD150 (SLAM), CD229 (Ly9) and CD244 (2B4)». *Tissue Antigens.* vol. 64, no 2, p. 132-144.
- Roof, D. J., A. Hayes, M. Adamian, A. H. Chishti et T. Li. 1997. «Molecular characterization of abLIM, a novel actin-binding and double zinc finger protein». *J Cell Biol.* vol. 138, no 3, p. 575-588.
- Rosenbauer, F., et D. G. Tenen. 2007. «Transcription factors in myeloid development: balancing differentiation with transformation». *Nat Rev Immunol.* vol. 7, no 2, p. 105-117.
- Rosmarin, A. G., S. C. Weil, G. L. Rosner, J. D. Griffin, M. A. Arnaout et D. G. Tenen. 1989. «Differential expression of CD11b/CD18 (Mo1) and myeloperoxidase genes during myeloid differentiation». *Blood.* vol. 73, no 1, p. 131-136.
- Rothenberg, E. V., et T. Taghon. 2005. «Molecular genetics of T cell development». *Annu Rev Immunol.* vol. 23, p. 601-649.
- Rothhammer, T., I. Poser, F. Soncin, F. Bataille, M. Moser et A. K. Bosserhoff. 2005. «Bone morphogenic proteins are overexpressed in malignant melanoma and promote cell invasion and migration». *Cancer Res.* vol. 65, no 2, p. 448-456.
- Rowley, P. T., B. A. Farley, S. LaBella, R. Giuliano et J. F. Leary. 1992. «Single K562 human leukemia cells express and are inducible for both erythroid and megakaryocytic antigens». *Int J Cell Cloning.* vol. 10, no 4, p. 232-240.
- Ru, M., C. Shustik et E. Rassart. 1993. «Graffi murine leukemia virus: molecular cloning and characterization of the myeloid leukemia-inducing agent». *J Virol.* vol. 67, no 8, p. 4722-4731.

- Rubio, M. T., N. Dhedin, C. Boucheix, J. H. Bourhis, O. Reman, J. M. Boiron, J. H. Gallo, V. Lheritier, X. Thomas, D. Fiere et J. P. Vernant. 2003. «Adult T-biphenotypic acute leukaemia: clinical and biological features and outcome». *Br J Haematol.* vol. 123, no 5, p. 842-849.
- Rutella, S., C. Rumi, P. Puggioni, T. Barberi, A. Di Mario, L. M. Larocca et G. Leone. 1999. «Expression of thrombospondin receptor (CD36) in B-cell chronic lymphocytic leukemia as an indicator of tumor cell dissemination». *Haematologica.* vol. 84, no 5, p. 419-424.
- Saino, M., T. Maruyama, T. Sekiya, T. Kayama et Y. Murakami. 2004. «Inhibition of angiogenesis in human glioma cell lines by antisense RNA from the soluble guanylate cyclase genes, GUCY1A3 and GUCY1B3». *Oncol Rep.* vol. 12, no 1, p. 47-52.
- Sambrook, J., et D. W. Russell. 2001. «Molecular cloning: a laboratory manual, 3rd ed.». Cold Spring Harbor Laboratory Press Cold Spring Harbor, N.Y.
- Samii, K., et E. Pasteur. 1998. «Images in hematology. Emperipolesis». *Am J Hematol.* vol. 59, no 1, p. 64.
- Sathyanarayana, P., M. P. Menon, O. Bogacheva, O. Bogachev, K. Niss, W. S. Kapelle, E. Houde, J. Fang et D. M. Wojchowski. 2007. «Erythropoietin modulation of podocalyxin and a proposed erythroblast niche». *Blood.* vol. 110, no 2, p. 509-518.
- Sato, T., A. Fuse, M. Eguchi, Y. Hayashi, R. Ryo, M. Adachi, Y. Kishimoto, M. Teramura, H. Mizoguchi, Y. Shima et et al. 1989. «Establishment of a human leukaemic cell line (CMK) with megakaryocytic characteristics from a Down's syndrome patient with acute megakaryoblastic leukaemia». *Br J Haematol.* vol. 72, no 2, p. 184-190.
- Sawyers, C. L., C. T. Denny et O. N. Witte. 1991. «Leukemia and the disruption of normal hematopoiesis». *Cell.* vol. 64, no 2, p. 337-350.
- Scamurra, D. O., F. R. Davey, D. A. Nelson, A. S. Kurec et J. Goldberg. 1983. «Acute leukemia presenting with myeloid and lymphoid cell markers». *Ann Clin Lab Sci.* vol. 13, no 6, p. 496-502.
- Schabath, H., S. Runz, S. Joumaa et P. Altevogt. 2006. «CD24 affects CXCR4 function in pre-B lymphocytes and breast carcinoma cells». *J Cell Sci.* vol. 119, no Pt 2, p. 314-325.
- Schmidt, M., D. A. Carbonaro, C. Speckmann, M. Wissler, J. Bohnsack, M. Elder, B. J. Aronow, J. A. Nolte, D. B. Kohn et C. von Kalle. 2003. «Clonality analysis after retroviral-mediated gene transfer to CD34+ cells from the cord blood of ADA-deficient SCID neonates». *Nat Med.* vol. 9, no 4, p. 463-468.
- Schuh, A. H., A. J. Tipping, A. J. Clark, I. Hamlett, B. Guyot, F. J. Iborra, P. Rodriguez, J. Strouboulis, T. Enver, P. Vyas et C. Porcher. 2005. «ETO-2 associates with SCL in erythroid cells and megakaryocytes and provides repressor functions in erythropoiesis». *Mol Cell Biol.* vol. 25, no 23, p. 10235-10250.

- Schwab, M., K. Alitalo, K. H. Klempnauer, H. E. Varmus, J. M. Bishop, F. Gilbert, G. Brodeur, M. Goldstein et J. Trent. 1983. «Amplified DNA with limited homology to myc cellular oncogene is shared by human neuroblastoma cell lines and a neuroblastoma tumour». *Nature*. vol. 305, no 5931, p. 245-248.
- Schwartz-Albiez, R., B. Dorken, W. Hofmann et G. Moldenhauer. 1988. «The B cell-associated CD37 antigen (gp40-52). Structure and subcellular expression of an extensively glycosylated glycoprotein». *J Immunol*. vol. 140, no 3, p. 905-914.
- Sempowski, G. D., D. M. Lee, R. E. Kaufman et B. F. Haynes. 1999. «Structure and function of the CD7 molecule». *Crit Rev Immunol*. vol. 19, no 4, p. 331-348.
- Senis, Y. A., W. Sangrar, R. A. Zirngibl, A. W. Craig, D. H. Lee et P. A. Greer. 2003. «Fps/Fes and Fer non-receptor protein-tyrosine kinases regulate collagen- and ADP-induced platelet aggregation». *J Thromb Haemost*. vol. 1, no 5, p. 1062-1070.
- Sepulveda, J. L., et C. Wu. 2006. «The parvins». *Cell Mol Life Sci*. vol. 63, no 1, p. 25-35.
- Shackelford, D., C. Kenific, A. Blusztajn, S. Waxman et R. Ren. 2006. «Targeted degradation of the AML1/MDS1/EVI1 oncoprotein by arsenic trioxide». *Cancer Res*. vol. 66, no 23, p. 11360-11369.
- Shafarenko, M., J. Mahler, C. Cochran, A. Kisielewski, E. Golding, R. Wiseman et T. Goodrow. 1997. «Similar incidence of K-ras mutations in lung carcinomas of FVB/N mice and FVB/N mice carrying a mutant p53 transgene». *Carcinogenesis*. vol. 18, no 7, p. 1423-1426.
- Shan, Z., T. Parker et J. S. Wiest. 2004. «Identifying novel homozygous deletions by microsatellite analysis and characterization of tumor suppressor candidate 1 gene, TUSC1, on chromosome 9p in human lung cancer». *Oncogene*. vol. 23, no 39, p. 6612-6620.
- Sheppard, F. R., M. R. Kelher, E. E. Moore, N. J. McLaughlin, A. Banerjee et C. C. Silliman. 2005. «Structural organization of the neutrophil NADPH oxidase: phosphorylation and translocation during priming and activation». *J Leukoc Biol*. vol. 78, no 5, p. 1025-1042.
- Shimizu, S., T. Nagasawa, O. Katoh, N. Komatsu, J. Yokota et K. Morishita. 2002. «EVI1 is expressed in megakaryocyte cell lineage and enforced expression of EVI1 in UT-7/GM cells induces megakaryocyte differentiation». *Biochem Biophys Res Commun*. vol. 292, no 3, p. 609-616.
- Shinnick, T. M., R. A. Lerner et J. G. Sutcliffe. 1981. «Nucleotide sequence of Moloney murine leukaemia virus». *Nature*. vol. 293, no 5833, p. 543-548.
- Shirihai, O. S., T. Gregory, C. Yu, S. H. Orkin et M. J. Weiss. 2000. «ABC-me: a novel mitochondrial transporter induced by GATA-1 during erythroid differentiation». *Embo J*. vol. 19, no 11, p. 2492-2502.

- Shivdasani, R. A. 2001. «Molecular and transcriptional regulation of megakaryocyte differentiation». *Stem Cells*. vol. 19, no 5, p. 397-407.
- Sijts, E. J., C. J. Leupers, E. A. Mengede, W. A. Loenen, P. J. van den Elsen et C. J. Melief. 1994. «Cloning of the MCF1233 murine leukemia virus and identification of sequences involved in viral tropism, oncogenicity and T cell epitope formation». *Virus Res*. vol. 34, no 3, p. 339-349.
- Silver, J., et C. E. Buckler. 1986. «A preferred region for integration of Friend murine leukemia virus in hematopoietic neoplasms is closely linked to the Int-2 oncogene». *J Virol*. vol. 60, no 3, p. 1156-1158.
- Slayton, W. B., A. Georgelas, L. J. Pierce, K. S. Elenitoba-Johnson, S. S. Perry, M. Marx et G. J. Spangrude. 2002. «The spleen is a major site of megakaryopoiesis following transplantation of murine hematopoietic stem cells». *Blood*. vol. 100, no 12, p. 3975-3982.
- Sobocka, M. B., T. Sobocki, A. Babinska, J. H. Hartwig, M. Li, Y. H. Ehrlich et E. Kornecki. 2004. «Signaling pathways of the F11 receptor (F11R; a.k.a. JAM-1, JAM-A) in human platelets: F11R dimerization, phosphorylation and complex formation with the integrin GPIIb/IIIa». *J Recept Signal Transduct Res*. vol. 24, no 1-2, p. 85-105.
- Souyri, M., I. Vigon, J. F. Penciolelli, J. M. Heard, P. Tambourin et F. Wendling. 1990. «A putative truncated cytokine receptor gene transduced by the myeloproliferative leukemia virus immortalizes hematopoietic progenitors». *Cell*. vol. 63, no 6, p. 1137-1147.
- Speck, N. A., et D. Baltimore. 1987. «Six distinct nuclear factors interact with the 75-base-pair repeat of the Moloney murine leukemia virus enhancer». *Mol Cell Biol*. vol. 7, no 3, p. 1101-1110.
- Sprague, D. L., J. M. Sowa, B. D. Elzey et T. L. Ratliff. 2007. «The role of platelet CD154 in the modulation in adaptive immunity». *Immunol Res*. vol. 39, no 1-3, p. 185-193.
- Starling, G. C., G. S. Whitney, A. W. Siadak, M. B. Llewellyn, M. A. Bowen, A. G. Farr et A. A. Aruffo. 1996. «Characterization of mouse CD6 with novel monoclonal antibodies which enhance the allogeneic mixed leukocyte reaction». *Eur J Immunol*. vol. 26, no 4, p. 738-746.
- Steindler, C., Z. Li, M. Algarte, A. Alcover, V. Libri, J. Ragimbeau et S. Pellegrini. 2004. «Jamip1 (marlin-1) defines a family of proteins interacting with janus kinases and microtubules». *J Biol Chem*. vol. 279, no 41, p. 43168-43177.
- Stephenson, J.R. 1980. «Molecular biology of RNA tumor viruses». Academic Press.
- Stevens, A., M. Bock, S. Ellis, P. LeTissier, K. N. Bishop, M. W. Yap, W. Taylor et J. P. Stoye. 2004. «Retroviral capsid determinants of Fv1 NB and NR tropism». *J Virol*. vol. 78, no 18, p. 9592-9598.

Stothard, P. 2000. «The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences». *Biotechniques*. vol. 28, no 6, p. 1102, 1104.

Stoye, J. P., et J. M. Coffin. 1987. «The four classes of endogenous murine leukemia virus: structural relationships and potential for recombination». *J Virol*. vol. 61, no 9, p. 2659-2669.

Strobl, H., C. Scheinecker, B. Csmarits, O. Majdic et W. Knapp. 1995. «Flow cytometric analysis of intracellular CD68 molecule expression in normal and malignant haemopoiesis». *Br J Haematol*. vol. 90, no 4, p. 774-782.

Stubbs, M. C., Y. M. Kim, A. V. Krivtsov, R. D. Wright, Z. Feng, J. Agarwal, A. L. Kung et S. A. Armstrong. 2007. «MLL-AF9 and FLT3 cooperation in acute myelogenous leukemia: development of a model for rapid therapeutic assessment». *Leukemia*.

Suzuki, T., K. Minehata, K. Akagi, N. A. Jenkins et N. G. Copeland. 2006. «Tumor suppressor gene identification using retroviral insertional mutagenesis in Blm-deficient mice». *Embo J*. vol. 25, no 14, p. 3422-3431.

T'Sas, F., C. Brenner, S. Mauen, P. Putmans, D. Monte, C. Van Lint, M. Moser, J. L. Baert et Y. de Launoit. 2005. «Expression of the Ets transcription factor Erm is regulated through a conventional PKC signaling pathway in the Molt4 lymphoblastic cell line». *FEBS Lett*. vol. 579, no 1, p. 66-70.

Tallman, M. S., D. Neuberg, J. M. Bennett, C. J. Francois, E. Paietta, P. H. Wiernik, G. Dewald, P. A. Cassileth, M. M. Oken et J. M. Rowe. 2000. «Acute megakaryocytic leukemia: the Eastern Cooperative Oncology Group experience». *Blood*. vol. 96, no 7, p. 2405-2411.

Tanaka, M., Y. Aze et T. Fujita. 1997. «Adhesion molecule LFA-1/ICAM-1 influences on LPS-induced megakaryocytic emperipoiesis in the rat bone marrow». *Vet Pathol*. vol. 34, no 5, p. 463-466. >.

Tani, T., J. Ylänne et I. Virtanen. 1996. «Expression of megakaryocytic and erythroid properties in human leukemic cells». *Exp Hematol*. vol. 24, no 2, p. 158-168.

Tassies, D., E. Montserrat, J. C. Reverter, N. Villamor, M. Rovira et C. Rozman. 1995. «Myelomonocytic antigens in B-cell chronic lymphocytic leukemia». *Leuk Res*. vol. 19, no 11, p. 841-848.

Tatusova, T. A., et T. L. Madden. 1999. «BLAST 2 Sequences, a new tool for comparing protein and nucleotide sequences». *FEMS Microbiol Lett*. vol. 174, no 2, p. 247-250.

Telen, M. J. 2005. «Erythrocyte adhesion receptors: blood group antigens and related molecules». *Transfus Med Rev*. vol. 19, no 1, p. 32-44.

Thomas, A., K. D. Gray et M. J. Roth. 1997. «Analysis of mutations within the cytoplasmic domain of the Moloney murine leukemia virus transmembrane protein». *Virology*. vol. 227, no 2, p. 305-313.

- Thompson, J. D., D. G. Higgins et T. J. Gibson. 1994. «CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice». *Nucleic Acids Res.* vol. 22, no 22, p. 4673-4680.
- Tipper, C. H., C. E. Bencsics et J. M. Coffin. 2005. «Characterization of hortulanus endogenous murine leukemia virus, an endogenous provirus that encodes an infectious murine leukemia virus of a novel subgroup». *J Virol.* vol. 79, no 13, p. 8316-8329.
- Tolmachova, T., M. Abrink, C. E. Futter, K. S. Authi et M. C. Seabra. 2007. «Rab27b regulates number and secretion of platelet dense granules». *Proc Natl Acad Sci U S A.* vol. 104, no 14, p. 5872-5877.
- Tomonaga, K., et J. M. Coffin. 1999. «Structures of endogenous noncotropic murine leukemia virus (MLV) long terminal repeats in wild mice: implication for evolution of MLVs». *J Virol.* vol. 73, no 5, p. 4327-4340.
- Trinder, D., et E. Baker. 2003. «Transferrin receptor 2: a new molecule in iron metabolism». *Int J Biochem Cell Biol.* vol. 35, no 3, p. 292-296.
- Tsang, A. P., Y. Fujiwara, D. B. Hom et S. H. Orkin. 1998. «Failure of megakaryopoiesis and arrested erythropoiesis in mice lacking the GATA-1 transcriptional cofactor FOG». *Genes Dev.* vol. 12, no 8, p. 1176-1188.
- Tsang, A. P., J. E. Visvader, C. A. Turner, Y. Fujiwara, C. Yu, M. J. Weiss, M. Crossley et S. H. Orkin. 1997. «FOG, a multitype zinc finger protein, acts as a cofactor for transcription factor GATA-1 in erythroid and megakaryocytic differentiation». *Cell.* vol. 90, no 1, p. 109-119.
- Tsiamis, A. C., P. Hayes, H. Box, A. H. Goodall, P. R. Bell et N. P. Brindle. 2000. «Characterization and regulation of the receptor tyrosine kinase Tie-1 in platelets». *J Vasc Res.* vol. 37, no 6, p. 437-442.
- Tsiftoglou, A. S., et W. Wong. 1985. «Molecular and cellular mechanisms of leukemic hemopoietic cell differentiation: an analysis of the Friend system». *Anticancer Res.* vol. 5, no 1, p. 81-99.
- Tsuchihara, K., V. Lapin, C. Bakal, H. Okada, L. Brown, M. Hirota-Tsuchihara, K. Zaugg, A. Ho, A. Itie-Youten, M. Harris-Brandts, R. Rottapel, C. D. Richardson, S. Benchimol et T. W. Mak. 2005. «Ckap2 regulates aneuploidy, cell cycling, and cell death in a p53-dependent manner». *Cancer Res.* vol. 65, no 15, p. 6685-6691.
- Tsujimura, K., Y. Obata, Y. Matsudaira, K. Nishida, Y. Akatsuka, Y. Ito, A. Demachi-Okamura, K. Kuzushima et T. Takahashi. 2006. «Characterization of murine CD160+ CD8+ T lymphocytes». *Immunol Lett.* vol. 106, no 1, p. 48-56.
- Tusher, V. G., R. Tibshirani et G. Chu. 2001. «Significance analysis of microarrays applied to the ionizing radiation response». *Proc Natl Acad Sci U S A.* vol. 98, no 9, p. 5116-5121.

- Underhill, D. M., et A. Ozinsky. 2002. «Phagocytosis of microbes: complexity in action». *Annu Rev Immunol.* vol. 20, p. 825-852.
- Uozumi, K., M. Otsuka, N. Ohno, T. Moriyama, S. Suzuki, S. Shimotakahara, I. Matsumura, S. Hanada et T. Arima. 2000. «Establishment and characterization of a new human megakaryoblastic cell line (SET-2) that spontaneously matures to megakaryocytes and produces platelet-like particles». *Leukemia.* vol. 14, no 1, p. 142-152.
- Utrera, R., L. Collavin, D. Lazarevic, D. Delia et C. Schneider. 1998. «A novel p53-inducible gene coding for a microtubule-localized protein with G2-phase-specific expression». *Embo J.* vol. 17, no 17, p. 5015-5025.
- Uzan, G., M. Prenant, M. H. Prandini, F. Martin et G. Marguerie. 1991. «Tissue-specific expression of the platelet GPIIb gene». *J Biol Chem.* vol. 266, no 14, p. 8932-8939.
- Vainchenker, W., N. Methia, N. Debili, M. Titeux et F. Wendling. 1995. «c-mpl, the thrombopoietin receptor». *Thromb Haemost.* vol. 74, no 1, p. 526-528.
- van Lohuizen, M., M. Breuer et A. Berns. 1989. «N-myc is frequently activated by proviral insertion in MuLV-induced T cell lymphomas». *Embo J.* vol. 8, no 1, p. 133-136.
- Vannucchi, A. M., S. Linari, C. Cellai, M. J. Koury et F. Paoletti. 1997. «Constitutive and inducible expression of megakaryocyte-specific genes in Friend erythroleukaemia cells». *Br J Haematol.* vol. 99, no 3, p. 500-508.
- Vannucchi, A. M., F. Paoletti, S. Linari, C. Cellai, R. Caporale, P. R. Ferrini, M. Sanchez, G. Migliaccio et A. R. Migliaccio. 2000. «Identification and characterization of a bipotent (erythroid and megakaryocytic) cell precursor from the spleen of phenylhydrazine-treated mice». *Blood.* vol. 95, no 8, p. 2559-2568.
- Vehvilainen, P., M. Hyytiainen et J. Keski-Oja. 2003. «Latent transforming growth factor-beta-binding protein 2 is an adhesion protein for melanoma cells». *J Biol Chem.* vol. 278, no 27, p. 24705-24713.
- Venesio, T., D. Taverna, N. E. Hynes, R. Deed, D. MacAllan, F. Ciardiello, E. M. Valverius, D. S. Salomon, R. Callahan et G. Merlo. 1992. «The int-2 gene product acts as a growth factor and substitutes for basic fibroblast growth factor in promoting the differentiation of a normal mouse mammary epithelial cell line». *Cell Growth Differ.* vol. 3, no 1, p. 63-71.
- Villemur, R., Y. Monczak, E. Rassart, C. Kozak et P. Jolicœur. 1987. «Identification of a new common provirus integration site in gross passage A murine leukemia virus-induced mouse thymoma DNA». *Mol Cell Biol.* vol. 7, no 1, p. 512-522.
- Visvader, J., et J. M. Adams. 1993. «Megakaryocytic differentiation induced in 416B myeloid cells by GATA-2 and GATA-3 transgenes or 5-azacytidine is tightly coupled to GATA-1 expression». *Blood.* vol. 82, no 5, p. 1493-1501.
- Voisin, V., C. Barat, T. Hoang et E. Rassart. 2006. «Novel insights into the pathogenesis of the Graffi murine leukemia retrovirus». *J Virol.* vol. 80, no 8, p. 4026-4037.

- Voisin, V., et E. Rassart. 2007. «Complete genome sequences of the two viral variants of the Graffi MuLV: Phylogenetic relationship with other murine leukemia retroviruses». *Virology*.
- Von Boros, j., et A. Korenyi. 1931. «Über einen fall von akuter megakaryocyblasten-leukämie, zugleich einige bemerkungen zum Problem der akuten leukemie.». *Z Klin Med*. vol. 118, p. 679-718.
- Voss, B., J. N. McLaughlin, M. Holinstat, R. Zent et H. E. Hamm. 2007. «PAR1, but not PAR4, activates human platelets through a Gi/o/phosphoinositide-3 kinase signaling axis». *Mol Pharmacol*. vol. 71, no 5, p. 1399-1406.
- Voytek, P., et C. A. Kozak. 1989. «Nucleotide sequence and mode of transmission of the wild mouse ecotropic virus, HoMuLV». *Virology*. vol. 173, no 1, p. 58-67.
- Wakioka, T., A. Sasaki, R. Kato, T. Shouda, A. Matsumoto, K. Miyoshi, M. Tsuneoka, S. Komiya, R. Baron et A. Yoshimura. 2001. «Spred is a Sprouty-related suppressor of Ras signalling». *Nature*. vol. 412, no 6847, p. 647-651.
- Wakkach A, Cottrez F, Groux H. 2001. «Differentiation of regulatory T cells 1 is induced by CD2 costimulation». *J Immunol*. vol. 167, no 6, p. 3107-3113.
- Wang, N., K. K. Lin, Z. Lu, K. S. Lam, R. Newton, X. Xu, Z. Yu, G. N. Gill et B. Andersen. 2007. «The LIM-only factor LMO4 regulates expression of the BMP7 gene through an HDAC2-dependent mechanism, and controls cell proliferation and apoptosis of mammary epithelial cells». *Oncogene*.
- Wang, P. L., S. O'Farrell, C. Clayberger et A. M. Krensky. 1992. «Identification and molecular cloning of tactile. A novel human T cell activation antigen that is a member of the Ig gene superfamily». *J Immunol*. vol. 148, no 8, p. 2600-2608.
- Wang, S., N. Nath, M. Adlam et S. Chellappan. 1999. «Prohibitin, a potential tumor suppressor, interacts with RB and regulates E2F function». *Oncogene*. vol. 18, no 23, p. 3501-3510.
- Wang, X., J. D. Crispino, D. L. Letting, M. Nakazawa, M. Poncz et G. A. Blobel. 2002. «Control of megakaryocyte-specific gene expression by GATA-1 and FOG-1: role of Ets transcription factors». *Embo J*. vol. 21, no 19, p. 5225-5234.
- Wang, Y., et R. A. Casero, Jr. 2006. «Mammalian polyamine catabolism: a therapeutic target, a pathological problem, or both?». *J Biochem (Tokyo)*. vol. 139, no 1, p. 17-25.
- Weiss, M. J., et S. H. Orkin. 1995. «GATA transcription factors: key regulators of hematopoiesis». *Exp Hematol*. vol. 23, no 2, p. 99-107.
- Welch, J. J., J. A. Watts, C. R. Vakoc, Y. Yao, H. Wang, R. C. Hardison, G. A. Blobel, L. A. Chodosh et M. J. Weiss. 2004. «Global regulation of erythroid gene expression by transcription factor GATA-1». *Blood*. vol. 104, no 10, p. 3136-3147.

- Weyrich, A. S., et G. A. Zimmerman. 2004. «Platelets: signaling cells in the immune continuum». *Trends Immunol.* vol. 25, no 9, p. 489-495.
- Whelan, S., P. Lio et N. Goldman. 2001. «Molecular phylogenetics: state-of-the-art methods for looking into the past». *Trends Genet.* vol. 17, no 5, p. 262-272.
- Wilson, B. D., M. Ii, K. W. Park, A. Suli, L. K. Sorensen, F. Larrieu-Lahargue, L. D. Urness, W. Suh, J. Asai, G. A. Kock, T. Thorne, M. Silver, K. R. Thomas, C. B. Chien, D. W. Losordo et D. Y. Li. 2006. «Netrins promote developmental and therapeutic angiogenesis». *Science.* vol. 313, no 5787, p. 640-644..
- Winau, F., V. Schwierzeck, R. Hurwitz, N. Rimmel, P. A. Sieling, R. L. Modlin, S. A. Porcelli, V. Brinkmann, M. Sugita, K. Sandhoff, S. H. Kaufmann et U. E. Schaible. 2004. «Saposin C is required for lipid presentation by human CD1b». *Nat Immunol.* vol. 5, no 2, p. 169-174. >.
- Witke, W., J. D. Sutherland, A. Sharpe, M. Arai et D. J. Kwiatkowski. 2001. «Profilin I is essential for cell survival and cell division in early mouse development». *Proc Natl Acad Sci U S A.* vol. 98, no 7, p. 3832-3836.
- Wolff, L., R. Koller, X. Hu et M. R. Anver. 2003. «A Moloney murine leukemia virus-based retrovirus with 4070A long terminal repeat sequences induces a high incidence of myeloid as well as lymphoid neoplasms». *J Virol.* vol. 77, no 8, p. 4965-4971.
- Wu, Z., et R. A. Irizarry. 2005. «Stochastic models inspired by hybridization theory for short oligonucleotide arrays». *J Comput Biol.* vol. 12, no 6, p. 882-893.
- Wunderlich, V., et H. Bielka. 2006. «Arnold Graffi (1910-2006) : A pioneer of experimental cancer research». *J Cancer Res Clin Oncol.*
- Xu, Z., S. Huang, L. S. Chang, A. D. Agulnick et S. J. Brandt. 2003. «Identification of a TAL1 target gene reveals a positive role for the LIM domain-binding protein Ldb1 in erythroid gene expression and differentiation». *Mol Cell Biol.* vol. 23, no 21, p. 7585-7599.
- Xue, H., D. O'Neill, J. Morrow et A. Bank. 1999. «A novel mouse gene, HemT, encoding an hematopoietic cell-specific transcript». *Gene.* vol. 231, no 1-2, p. 49-58.
- Yan, Y., V. Narayanan et C. Lagenaur. 1996. «Expression of members of the proteolipid protein gene family in the developing murine central nervous system». *J Comp Neurol.* vol. 370, no 4, p. 465-478.
- Yancopoulos, G. D., P. D. Nisen, A. Tesfaye, N. E. Kohl, M. P. Goldfarb et F. W. Alt. 1985. «N-myc can cooperate with ras to transform normal cells in culture». *Proc Natl Acad Sci U S A.* vol. 82, no 16, p. 5455-5459.
- Yang, Y., M. Marcello, V. Endris, R. Saffrich, R. Fischer, M. F. Trendelenburg, R. Sprengel et G. Rappold. 2006. «MEGAP impedes cell migration via regulating actin and microtubule dynamics and focal complex formation». *Exp Cell Res.* vol. 312, no 12, p. 2379-2393.

- Yates, K. E., et J. C. Gasson. 1996. «Role of c-Fes in normal and neoplastic hematopoiesis». *Stem Cells*. vol. 14, no 1, p. 117-123.
- Ye, M., et T. Graf. 2007. «Early decisions in lymphoid development». *Curr Opin Immunol*. vol. 19, no 2, p. 123-128.
- Yeh, C. T., S. C. Lu, I. C. Tseng, H. Y. Lai, M. L. Tsao, S. F. Huang et Y. F. Liaw. 2003. «Antisense overexpression of BMAL2 enhances cell proliferation». *Oncogene*. vol. 22, no 34, p. 5306-5314.
- Yeoh, E. J., M. E. Ross, S. A. Shurtleff, W. K. Williams, D. Patel, R. Mahfouz, F. G. Behm, S. C. Raimondi, M. V. Relling, A. Patel, C. Cheng, D. Campana, D. Wilkins, X. Zhou, J. Li, H. Liu, C. H. Pui, W. E. Evans, C. Naeve, L. Wong et J. R. Downing. 2002. «Classification, subtype discovery, and prediction of outcome in pediatric acute lymphoblastic leukemia by gene expression profiling». *Cancer Cell*. vol. 1, no 2, p. 133-143.
- Yin, T., et L. Li. 2006. «The stem cell niches in bone». *J Clin Invest*. vol. 116, no 5, p. 1195-1201.
- Yizhar, O., U. Matti, R. Melamed, Y. Hagalili, D. Bruns, J. Rettig et U. Ashery. 2004. «Tomosyn inhibits priming of large dense-core vesicles in a calcium-dependent manner». *Proc Natl Acad Sci U S A*. vol. 101, no 8, p. 2578-2583.
- Youssefian, T., A. Drouin, J. M. Masse, J. Guichard et E. M. Cramer. 2002. «Host defense role of platelets: engulfment of HIV and *Staphylococcus aureus* occurs in a specific subcellular compartment and is enhanced by platelet activation». *Blood*. vol. 99, no 11, p. 4021-4029.
- Yuen, P. H., Y. H. Khang et C. Matherne. 1992. «Virulence in lymphomagenesis is modulated by alterations in proteins binding to CORE and GRE-LVa elements of the MoMuLV enhancer». *Leukemia*. vol. 6 Suppl 3, p. 76S-82S.
- Zenkhusen, J. C., C. J. Conti et E. D. Green. 2001. «Mutational and functional analyses reveal that ST7 is a highly conserved tumor-suppressor gene on human chromosome 7q31». *Nat Genet*. vol. 27, no 4, p. 392-398.
- Zeuner, A., M. Signore, D. Martinetti, M. Bartucci, C. Peschle et R. De Maria. 2007. «Chemotherapy-induced thrombocytopenia derives from the selective death of megakaryocyte progenitors and can be rescued by stem cell factor». *Cancer Res*. vol. 67, no 10, p. 4767-4773.
- Zhang, C., M. A. Thornton, M. A. Kowalska, B. S. Sachis, M. Feldman, M. Poncz, S. E. McKenzie et M. P. Reilly. 2001. «Localization of distal regulatory domains in the megakaryocyte-specific platelet basic protein/platelet factor 4 gene locus». *Blood*. vol. 98, no 3, p. 610-617.

Appendice A

		T1	T2	T3	B1	B2	B3	M	E1	E2	E3	Mk1	Mk2	Mk3	
1416330	at	Cd81	0,3	-0,6	0,0	1,0	0,7	1,6	-0,5	-0,1	-1,3	-1,3	-0,5	-0,5	1,3
1417268	at	Cd14	-1,1	-0,5	-1,4	-1,1	-1,2	-1,1	3,7	-0,4	-0,6	-0,8	1,0	1,2	2,2
1437670	x_at	Cd151	-1,5	-1,2	-0,7	-0,8	-1,0	-0,6	-0,3	1,5	0,2	-0,3	1,8	0,8	2,1
1420396	at	Cd160	1,6	1,7	2,9	-1,0	-1,0	1,9	-1,1	-1,0	-0,8	-0,9	-0,4	-1,1	-0,6
1419144	at	Cd163	-1,1	-1,1	-0,7	-0,8	-0,9	-1,0	6,0	-0,8	-0,7	-0,5	-0,7	0,5	1,9
1416440	at	Cd164	0,7	1,1	0,5	-0,3	-0,1	0,0	0,5	-0,5	-0,4	-0,6	-0,7	-0,4	0,1
1450570	a_at	Cd19	-2,7	-3,1	-2,0	5,1	4,9	4,9	-2,2	-1,1	-2,5	-2,4	-2,1	1,0	2,2
1449130	at	Cd1d1	0,0	0,3	-1,3	1,0	0,6	1,1	1,6	-0,4	-1,5	-1,9	0,0	-0,2	0,5
1418770	at	Cd2	3,3	4,1	2,6	0,4	1,2	2,1	-2,8	-2,9	-2,9	-3,0	-1,9	-0,4	0,2
1417224	a_at	Cd2bp2	1,1	0,9	0,8	-0,3	0,1	-0,2	-0,7	-0,2	0,7	0,4	-0,5	-0,9	-1,0
1448788	at	Cd200	-0,2	2,7	0,9	-0,1	-0,4	-0,3	-0,6	-0,5	-0,2	-0,9	-0,9	0,2	0,2
1419769	at	Cd22	-1,6	-2,1	-1,5	3,6	2,7	2,9	-1,1	-1,4	-1,7	-2,2	-1,5	1,8	2,1
1449991	at	Cd244	-2,1	-1,8	-1,7	0,4	2,9	3,8	1,3	-1,8	-1,7	-1,8	-1,0	2,1	1,4
1416034	at	Cd24a	-0,3	0,2	-1,8	0,3	0,0	0,1	-2,1	1,2	1,5	1,5	-1,1	0,6	0,0
1417597	at	Cd28	5,4	4,6	5,6	-1,5	-2,0	-2,1	-1,4	-2,0	-1,9	-1,9	-1,9	0,7	-1,5
1420907	at	Cd2ap	0,7	1,4	-0,3	0,5	0,9	0,8	0,6	-0,7	-1,6	-0,5	-1,4	-0,8	0,3
1422828	at	Cd3d	5,8	5,8	5,7	-2,5	-2,3	-2,4	-0,2	-1,7	-2,1	-2,1	-0,9	-2,1	-1,1
1422105	at	Cd3e	3,2	3,4	3,7	-1,0	-1,2	-1,0	-1,0	-1,1	-1,1	-1,3	-0,5	-1,2	-1,0
1419178	at	Cd3g	5,8	6,0	5,8	-2,4	-2,5	-2,5	-1,6	-2,1	-2,1	0,4	-1,8	-2,1	-0,8
1452539	a_at	Cd3z	3,4	3,4	3,7	-1,1	-1,1	-1,2	-1,0	-1,2	-0,8	-1,0	-0,9	-1,0	-1,3
1450513	at	Cd33	-0,6	-0,2	-0,3	-0,3	-0,3	-0,1	2,4	-0,3	0,0	-0,3	-0,3	-0,2	0,4
1416072	at	Cd34	-1,2	-1,1	-0,9	-0,9	-1,0	-1,0	4,6	-1,0	-0,6	-1,4	-0,5	5,5	-0,5
1450883	a_at	Cd36	-2,3	-2,2	-2,6	3,5	4,3	4,3	-1,6	2,1	1,1	-0,6	-2,4	-2,2	-1,3
1419206	at	Cd37	-1,2	-1,7	0,1	2,2	1,9	2,4	0,5	-0,8	-1,8	-1,1	-0,6	-0,4	0,6
1425736	at	Cd37	-0,5	-1,0	-0,9	2,4	2,3	1,9	0,0	-0,9	-1,8	-1,8	-0,6	-0,2	1,0
1433741	at	Cd38	-2,3	-2,5	0,2	3,2	4,1	3,3	-0,7	-1,3	-1,5	-2,5	-2,4	0,7	1,8
1450136	at	Cd38	-1,9	-1,7	0,2	3,0	3,9	3,1	-1,4	-1,4	-1,6	-2,0	-1,7	0,0	1,6
1427779	a_at	Cd4	1,0	-0,2	3,4	-0,4	-0,5	-0,5	-0,4	-0,3	-0,5	-0,2	-0,5	-0,5	-0,2
1423760	at	Cd44	-4,2	-0,7	0,2	0,0	-4,1	-2,0	2,1	1,2	1,9	1,9	1,2	0,9	1,7
1419554	at	Cd47	-1,6	-0,8	-2,4	-0,2	-0,6	0,0	0,3	1,3	0,4	0,7	0,9	0,9	1,1
1427301	at	Cd48	-5,9	-0,8	0,1	0,4	0,5	1,1	1,0	1,3	0,9	1,3	0,8	-1,4	0,5
1418353	at	Cd5	3,8	2,7	4,7	2,5	-0,5	2,3	-1,9	-2,3	-2,6	-2,3	-2,3	-2,2	-1,7
1449193	at	Cd51	-1,4	-1,7	-1,5	-1,8	-1,7	-1,9	0,6	2,7	2,2	0,8	-1,3	1,2	3,9
1460218	at		1,0	-1,7	2,4	2,6	3,0	2,4	1,8	-2,5	-3,9	-4,9	-1,7	0,5	1,2
1448617	at	Cd53	1,3	0,7	-1,2	2,1	2,0	2,0	2,3	-2,3	-3,4	-4,5	-0,2	0,4	0,8
1459744	at	Cd53	0,1	0,3	-1,1	1,5	1,7	1,4	1,0	-0,8	-1,1	-0,9	-1,1	-0,7	-0,5
1418710	at	Cd59a	-1,6	-1,4	-1,3	-1,7	-1,8	-1,6	-1,4	3,8	3,3	3,8	-0,8	-1,1	1,7
1429830	a_at	Cd59a	-1,9	-1,8	-2,1	-2,0	-1,6	-1,7	-1,9	4,5	4,0	4,5	-0,8	-1,2	2,0
1451910	a_at	Cd6	3,7	3,1	4,3	-0,7	-0,8	-0,9	-1,2	-1,5	-1,4	-0,9	-0,8	-1,6	-1,4
1449164	at	Cd68	-1,3	-1,5	-1,6	-2,4	-2,5	-2,3	3,4	0,1	-0,2	-1,3	3,2	3,5	3,0
1428735	at	Cd69	1,3	0,5	-0,5	0,9	0,7	1,5	0,9	-0,1	-1,7	-1,8	-0,1	-1,6	0,1
1419711	at	Cd7	0,7	1,7	-2,4	-2,3	-2,4	-2,1	-0,9	3,7	2,2	0,7	0,7	-1,8	2,2
1418830	at	Cd79a	-3,1	-3,0	-2,4	5,3	5,0	5,3	-2,2	-1,1	-2,8	-3,2	-2,5	2,1	2,6
1417640	at	Cd79b	-3,4	-3,7	-2,9	4,3	4,4	4,3	0,1	-1,0	-3,1	-3,6	-0,3	2,3	2,5
1425335	at	Cd8a	3,1	5,6	-0,8	-0,7	-0,8	-1,0	-0,7	-0,6	-0,8	-0,9	-0,6	-1,0	-0,8
1444078	at	Cd8a	3,7	5,8	-0,9	-1,1	-1,1	-0,6	-0,6	-0,9	-1,0	-0,8	-0,6	-1,0	-0,9
1451673	at	Cd8a	4,4	6,3	-1,1	-0,6	-0,7	-1,2	-1,1	-1,0	-1,2	-1,2	-0,7	-1,1	-0,8
1426170	a_at	Cd8b1	4,7	5,5	0,8	-1,0	-1,3	-1,4	-0,9	-1,2	-1,0	-1,1	-1,0	-1,0	-1,1
1448569	at	Cd8b1	-0,2	-0,4	0,3	0,7	0,5	0,9	0,1	-0,2	-1,4	-1,6	0,5	0,5	0,2
1416111	at	Cd83	-1,1	-1,1	-0,2	0,2	0,4	-0,3	0,7	-0,6	-0,7	-1,3	-1,1	2,2	3,0
1422875	at	Cd84	-0,1	-0,3	1,2	0,0	-0,2	-0,4	0,0	1,1	-1,8	-0,6	-0,6	0,9	0,8
1420404	at	Cd86	-0,7	-0,8	-0,7	-0,3	-0,6	1,0	0,4	-0,2	-0,2	-0,6	-0,6	0,9	2,4
1416066	at	Cd9	1,1	-2,5	2,5	-3,5	-3,6	-3,6	0,6	1,9	-1,8	-2,1	3,0	4,4	3,5
1419226	at	Cd96	0,9	2,7	1,9	-2,7	-2,3	-2,3	-1,0	0,3	-2,3	-2,7	1,9	3,4	2,1

Appendice B

Gènes spécifiques de Mk1, Mk2, Mk3																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																													</
------------------------------------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	----

Appendice B suite

Gene Symbol	Gene Title	ordre probesets	probesets redondants	ordre des gènes	% de homologie (Hs) *	T1_cal	T2_cal	T3_cal	B1_cal	B2_cal	B3_cal	M_cal	E1_cal	E2_cal	E3_cal	Mk1_cal	Mk2_cal	Mk3_cal
Fkbp9	FK506 binding protein 9	1	1/2	1	95	A	A	A	A	A	A	A	A	A	A	P	P	P
Fcer1a	Fc receptor, IgE, high affinity 1, alpha polypeptide	2		2	52	A	M	A	A	A	A	A	A	A	A	P	P	P
Sdpr	serum deprivation response	3	1/3	3	89	A	P	A	A	A	A	P	P	P	P	P	P	P
st8sia6	ST8 alpha-N-acetyl-neuraminide alpha-2,8-sialyltr	4	1/2	4	87	A	A	P	P	P	P	P	P	M	A	P	P	P
Sdpr	serum deprivation response	5	2/3		83	A	P	P	A	A	A	P	P	P	A	P	P	P
Gucy1a3	guanylate cyclase 1, soluble, alpha 3	6	1/3	5	93	A	A	A	A	A	A	P	A	A	A	P	P	P
700021P22R	RIKEN cDNA 4921540A03 gene	7		6	?	A	A	A	A	A	A	A	A	A	A	A	P	M
Meis1	Myeloid ecotropic viral integration site 1	8	1/4	7	82	A	P	A	A	A	A	A	A	A	A	P	P	P
St8sia6	Transcribed locus	9	2/2		87	A	A	P	A	A	A	A	A	A	A	P	P	P
Mpl	myeloproliferative leukemia virus oncogene	10	1/2	8	80	A	M	A	A	A	A	A	P	A	A	P	P	P
Nmyc1	neuroblastoma myc-related oncogene 1	11		9	85	A	M	A	P	P	P	P	P	A	A	P	P	P
Ica1	islet cell autoantigen 1	12	1/2	10	89	A	A	M	P	P	P	P	P	P	A	P	P	P
Slc35d3	solute carrier family 35, member D3	13		11	26	A	P	A	A	A	A	A	P	A	A	P	P	P
Fgf3	fibroblast growth factor 3	14	1/2	12	89	A	A	A	A	A	A	A	A	A	A	P	P	P
Cuedc1	CUE domain containing 1	15	1/3	13	86	A	A	A	A	A	A	A	A	A	A	A	P	P
Gulp1	GULP, engulfment adaptor PTB domain containin	16	1/3	14	92	A	A	M	A	A	A	A	A	A	A	P	P	P
Cuedc1	CUE domain containing 1	17	2/3		86	P	P	P	A	P	P	A	P	A	P	P	P	P
Ffar2	free fatty acid receptor 2	18	1/2	15	89	P	M	A	A	A	A	P	A	A	A	P	P	P
C3ar1	complement component 3a receptor 1	19	1/2	16	74	A	A	A	A	A	A	A	A	A	A	P	P	P
Vezf1	vascular endothelial zinc finger 1	20	1/2	17	98	P	P	P	P	P	P	P	P	P	P	P	P	P
Mfap3l	microfibrillar-associated protein 3-like	21		18	95	A	A	A	A	A	A	P	A	A	A	M	A	P
Zfp521	zinc finger protein 521(Evi3)	22	1/2	19	96	A	A	A	A	A	A	A	A	A	A	P	P	P
Tek	endothelial-specific receptor tyrosine kinase	23		20	95	P	A	A	A	A	A	A	P	P	A	P	P	P
Ifitm1	interferon induced transmembrane protein 1	24		21	79	A	P	P	A	P	P	P	P	P	A	P	P	P
610034B18R	RIKEN cDNA 2610034B18 gene	25		22	90	A	A	A	A	A	A	A	A	A	A	A	A	A
Ttpa	tocopherol (alpha) transfer protein	26		23	89	A	A	A	A	A	A	A	A	A	A	P	P	P
930519L02R	RIKEN cDNA 4930519L02 gene	27		24	?	A	P	A	A	A	P	P	A	P	P	P	P	P
Meis1	myeloid ecotropic viral integration site 1	28	2/4		82	A	A	A	A	A	A	A	A	A	A	P	P	P
Selp	selectin, platelet	29	1/2	25	64	P	P	A	A	A	A	P	P	A	A	P	P	P
Rnf180	ring finger protein 180	30		26	81	A	A	A	A	A	A	P	A	A	A	P	P	P
Cuedc1	CUE domain containing 1	31	3/3		86	P	A	P	A	A	A	P	A	A	A	P	P	P
Angpt1	Angiopoietin 1	32	1/4	27	99	P	P	A	A	A	A	P	A	A	P	P	P	P
Srgap3	SLIT-ROBO Rho GTPase activating protein 3	33	1/2	28	93	P	A	A	A	A	A	P	A	A	A	P	P	P
Syt11	synaptotagmin 11	34		29	98	M	A	P	A	A	P	A	A	A	A	A	A	A
Angpt1	angiopoietin 1	35	2/4		99	A	M	A	P	A	A	P	A	A	A	P	P	P
Gulp1	GULP, engulfment adaptor PTB domain containin	36	2/3		92	A	A	A	A	A	A	A	A	A	A	P	P	P
Meis1	myeloid ecotropic viral integration site 1	37	3/4		82	A	P	A	A	A	A	A	A	A	A	P	P	P
Srgap3	SLIT-ROBO Rho GTPase activating protein 3	38	2/2	30	93	A	A	A	A	A	A	A	A	A	A	A	A	A
Kit	Kit oncogene	39	1/3	31	83	A	A	A	A	P	P	A	P	P	P	P	P	P
Inpp5a	inositol polyphosphate-5-phosphatase A	40		32	98	A	P	A	P	P	P	P	P	A	A	P	P	P
Ffar2	G protein-coupled receptor 43	41	2/2		84	P	A	A	A	A	P	A	A	A	A	P	P	P
St8sia6	Transcribed locus	42	2/2	33	87	A	A	A	P	A	A	A	A	A	A	P	P	P
Cbfa2t3h	Core-binding factor, runt domain, alpha subunit 2,	43	1/2	34	?	A	A	A	P	A	P	P	P	A	A	P	P	P
Lhfp12	lipoma HMGIC fusion partner-like 2	44	1/3	35	91	A	A	A	A	A	A	P	A	A	A	P	P	P
Tmtc2	transmembrane and tetratricopeptide repeat contain	45		36	97	A	A	A	A	A	A	A	A	A	A	A	P	P
Sdpr	serum deprivation response	46	3/3		83	A	P	A	A	A	P	P	P	M	A	P	P	P
Gucy1a3	guanylate cyclase 1, soluble, alpha 3	47	2/3		93	A	A	A	A	A	A	A	A	A	A	P	P	P
Tpk1	thiamin pyrophosphokinase	48		37	91	P	P	P	P	P	P	P	A	A	A	P	P	P

Appendice C

Gènes spécifiques de Mk1, Mk2, Mk3, E1																					
ordre des probesets	F1	log (fold change)	SAM d-score	P-value	FDR	T1	T2	T3	B1	B2	B3	M	E1	E2	E3	Mk1	Mk2	Mk3	E1-E	Gene Symbol	lignée Mk
1	1437150 at	4.2	29.8	<0.001	0.00	-1.4	-1.0	-1.4	-1.5	-1.4	-1.4	-1.4	3.1	-1.0	-1.3	3.0	2.5	3.2	3.5	1700012H17Rik	
2	1419720 at	4.5	19.5	<0.001	0.01	-1.6	-1.5	-1.7	-1.6	-1.9	-1.9	-1.0	3.0	-0.9	-0.4	3.1	3.0	3.5	1.3	Gp9	Mk
3	1427169 at	1.9	15.4	<0.001	0.02	-0.6	-0.8	-0.5	-0.5	-0.5	-0.4	-0.8	1.2	-0.4	-0.6	1.0	1.2	1.6	1.5	Pard3b	
4	1438306 at	4.3	14.5	<0.001	0.03	-1.4	-1.2	-1.3	-1.3	-1.5	-1.4	-1.2	2.1	-1.1	-1.3	3.2	3.1	3.4	3.2	Rnf180	
5	1424923 at	6.6	13.1	<0.001	0.03	-2.6	-1.7	-3.1	-2.5	-2.4	-0.2	0.4	5.1	-2.1	-1.2	4.6	4.4	4.4	5.7	Serpina3g	
6	1460356 at	2.8	12.6	<0.001	0.03	-1.3	-1.0	-0.9	-1.0	-0.9	-1.1	-0.6	1.5	0.2	-0.7	1.5	2.2	2.1	0.7	Esam1	Mk
7	1457359 at	2.2	12.4	<0.001	0.03	-0.3	-0.7	-0.7	-0.8	-0.4	-0.2	-1.1	1.1	-0.9	-0.9	1.5	1.6	1.8	2.1	Inpp4b	
8	1430301 at	2.0	12.3	<0.001	0.03	-1.0	-0.9	-1.0	-0.3	-0.6	-0.4	-0.4	1.3	-0.7	0.0	1.1	1.8	1.2	1.8	Stxbp5	
9	1426782 at	4.5	12.1	<0.001	0.03	-1.7	-1.6	-1.6	-1.6	-1.5	-1.6	-1.4	4.1	-1.3	-0.3	2.6	3.2	2.6	3.7	Gpr125	
10	1417758 at	5.2	10.6	<0.001	0.04	-2.7	-1.6	-2.6	-2.9	-2.6	-2.7	-1.6	3.8	0.8	1.1	3.6	3.6	3.8	3.5	Itga2b	Mk
11	1447337 at	3.3	10.6	<0.001	0.04	-1.9	-0.8	-2.1	-0.9	-1.5	-1.8	-0.7	2.1	0.0	0.2	2.1	2.7	2.5	4.9	Dapp1	
12	1417689 a at	5.8	10.3	<0.001	0.04	-2.7	-1.7	-2.6	-2.9	-2.7	-2.9	-1.8	4.3	0.3	0.2	3.4	5.3	3.7	0.1	Pdzk1ip1	
13	1429601 x at	1.9	10.2	<0.001	0.04	-0.9	-1.1	-0.6	-0.6	-0.2	-0.2	-0.7	1.3	-0.6	0.2	1.6	0.7	1.3	1.9	1110019K23Rik	
14	1442833 at	0.6	9.6	<0.001	0.05	-0.2	-0.6	-0.3	0.0	-0.3	-0.1	-0.1	0.4	0.1	-0.1	0.4	0.4	0.5	-1.0	Ank	
15	1441425 at	3.8	9.6	<0.001	0.05	-1.8	-1.6	-1.6	-1.4	-1.5	-1.2	-1.4	2.6	-0.4	0.2	1.9	3.5	2.7	4.2	—	
16	1455477 s at	4.6	9.5	<0.001	0.05	-2.0	-1.6	-1.7	-2.1	-2.0	-1.9	-1.2	3.1	-0.2	-0.2	2.8	4.4	2.6	0.4	Pdzk1ip1	
17	1424894 at	0.9	9.4	<0.001	0.05	-0.3	-0.7	-0.2	-0.5	-0.4	-0.4	-0.1	0.6	0.0	-0.1	0.8	0.7	0.6	-0.2	Rab13	
18	1416827 at	3.7	9.2	<0.001	0.05	-2.0	-1.3	-1.2	-1.6	-1.7	-1.8	0.4	1.9	-0.6	-0.4	2.1	3.1	3.0	1.1	Tbxas1	Mk?
19	1429599 a at	2.5	8.8	<0.001	0.06	-1.4	-1.5	-1.3	-0.9	-0.5	0.1	-1.3	1.7	-0.6	0.4	1.7	1.4	2.1	2.7	1110019K23Rik	
20	1460735 at	1.5	8.7	<0.001	0.06	0.0	-0.4	0.0	-1.1	-0.7	-0.6	-0.6	1.2	-0.6	-0.2	1.3	0.7	0.9	1.5	Svil	
21	1434472 at	2.1	8.4	<0.001	0.06	-1.3	-1.4	-0.7	-0.4	-0.5	-0.7	0.0	2.0	-0.6	-0.5	1.1	1.3	1.6	1.8	Dusp3	
22	1439066 at	5.8	8.4	<0.001	0.06	-1.9	0.2	-3.1	-1.2	-2.4	-1.6	-2.2	2.3	-3.1	-0.7	4.3	5.2	4.2	3.2	Angpt1	Mk
23	1418788 at	3.4	8.3	<0.001	0.06	-1.0	-1.2	-1.4	-1.0	-1.2	-0.7	-1.0	1.1	-0.7	-1.0	3.0	2.4	2.6	2.0	Tek	
24	1445603 at	0.7	8.2	<0.001	0.06	-0.3	-0.2	-0.4	0.0	-0.3	-0.5	-0.1	0.5	-0.1	-0.2	0.7	0.4	0.4	0.6	Sri	
25	1448995 at	6.2	8.2	<0.001	0.06	-3.1	-0.9	-3.6	-3.7	-4.0	-3.4	1.1	4.8	-1.7	1.9	3.9	4.0	4.7	2.7	Cxcl4	Mk
26	1415984 at	1.2	8.1	<0.001	0.06	-0.4	-0.5	-0.5	-0.8	-0.7	-0.8	0.2	1.1	0.2	-0.2	0.7	0.8	0.8	1.8	Acadm	
27	1419518 at	3.1	8.1	<0.001	0.06	-2.0	-1.8	-0.2	-1.3	-1.7	-1.3	-0.1	2.2	-1.4	0.9	2.3	1.6	2.8	2.3	Tuba8	
28	1432184 a at	2.1	8.0	<0.001	0.06	-0.3	-1.1	-0.6	-1.4	-0.9	-0.7	-0.5	2.1	-0.4	0.0	1.2	1.2	1.4	2.6	2610204M08Rik	
29	1440045 at	0.8	7.9	<0.001	0.06	-0.5	-0.5	-0.6	0.2	-0.2	-0.2	-0.1	0.4	-0.1	0.1	0.5	0.5	0.6	0.4	Cryz1l	
30	1436216 s at	3.4	7.8	<0.001	0.06	-0.4	-2.1	-0.8	-2.4	-1.9	-1.1	-0.6	3.1	-0.5	0.4	1.8	1.9	2.4	3.8	2610204M08Rik	
31	1460626 at	0.9	7.8	<0.001	0.06	-0.2	-0.4	0.4	-0.4	-0.7	-0.5	-0.3	0.6	-0.3	-0.2	0.7	0.4	0.8	1.2	sept-11	
32	1452352 at	5.3	7.7	<0.001	0.06	-2.7	-2.4	0.5	-2.6	-1.5	-1.6	-1.2	5.0	-1.1	-1.9	2.4	2.7	4.4	4.8	Ctla2b	
33	1421441 at	5.1	7.7	<0.001	0.06	-1.6	-0.7	-2.0	-1.2	-1.9	-1.5	-1.8	1.7	-1.9	-1.0	3.6	4.6	3.6	3.3	Angpt1	Mk
34	1437670 x at	2.1	7.7	<0.001	0.06	-1.5	-1.2	-0.7	-0.8	-1.0	-0.6	-0.3	1.5	0.2	-0.3	1.8	0.8	2.1	1.5	Cd151	Mk
35	1456085 x at	2.2	7.6	<0.001	0.06	-1.2	-1.1	-0.7	-0.7	-0.9	-0.6	-0.3	1.4	-0.2	-0.3	1.8	0.8	2.1	1.5	Cd151	Mk
36	1417214 at	5.8	7.5	<0.001	0.06	-3.1	-2.5	-3.0	-3.0	-2.6	-3.1	-2.4	3.4	1.8	1.3	3.9	5.5	3.8	4.4	Rab27b	Mk
37	1424786 s at	0.8	7.5	<0.001	0.06	-0.1	-0.4	0.0	-0.4	-0.6	-0.4	-0.1	0.4	0.0	-0.2	0.5	0.8	0.6	0.3	Wdrx1	
38	1451232 at	1.9	7.4	<0.001	0.07	-1.4	-1.0	-0.7	-0.7	-0.9	-0.7	-0.2	1.3	0.1	-0.1	1.8	0.7	1.6	1.3	Cd151	Mk
39	1421095 a at	0.6	7.3	<0.001	0.07	-0.5	0.2	-0.2	-0.4	-0.3	-0.1	-0.4	0.3	-0.2	0.2	0.4	0.5	0.5	0.3	Trpc1	
40	1417605 s at	2.2	7.3	<0.001	0.07	-1.1	-0.9	-0.7	-1.2	-1.1	-1.5	1.7	1.6	-0.5	-0.8	1.3	1.6	1.6	0.9	Camk1	
41	1452072 at	4.2	7.2	<0.001	0.07	-2.0	-1.7	-1.7	-1.6	-1.5	-1.6	-1.4	2.0	-1.1	0.6	3.0	4.5	2.5	2.1	Myc1	
42	1451038 at	0.2	7.2	<0.001	0.07	-0.2	-0.1	0.0	0.0	-0.1	-0.1	0.0	0.1	-0.1	0.0	0.1	0.1	0.3	-0.1	Apln	
43	1446941 at	0.4	7.2	<0.001	0.07	-0.1	0.1	-0.2	-0.1	-0.1	-0.1	0.1	0.3	-0.4	-0.2	0.3	0.3	0.2	0.3	Sik	
44	1448402 at	1.7	7.2	<0.001	0.07	-0.9	-1.2	-0.8	-0.6	-1.1	-0.8	0.5	1.4	0.1	0.3	0.8	1.1	1.2	1.4	Tln1	
45	1448152 at	0.4	7.2	<0.001	0.07	0.0	-0.3	-0.2	-0.1	-0.1	-0.3	0.0	0.2	0.0	-0.1	0.4	0.0	0.3	-0.2	Igf2	
46	1431163 at	0.6	7.1	<0.001	0.07	-0.2	-0.3	-0.2	-0.1	-0.2	-0.1	0.1	0.2	-0.1	-0.4	0.6	0.4	0.5	0.1	2700046G09Rik	
47	1422445 at	3.1	7.1	<0.001	0.07	0.2	0.5	-0.2	-2.2	-2.7	-1.0	-0.9	2.1	-2.6	0.1	1.8	2.5	2.4	2.6	Itga6	
48	1442956 at	0.5	7.0	<0.001	0.07	-0.2	0.0	-0.3	0.0	-0.1	-0.2	-0.1	0.3	-0.4	0.0	0.3	0.4	0.3	0.3	Ppp1r13b	
49	1454973 at	0.9	7.0	<0.001	0.07	-0.4	-0.9	-0.1	-0.2	-0.2	0.2	-0.1	0.5	-0.6	-0.1	0.5	0.9	0.5	0.2	Atf7ip	
50	1421209 s at	0.5	7.0	<0.001	0.1	-0.1	-0.2	0.1	-0.1	-0.2	-0.2	-0.3	0.3	-0.1	-0.2	0.4	0.2	0.4	0.4	Ikbkg	

Appendice C suite

Gene Symbol		Gene Title	ordre des probesets	probesets redondants	ordre des gènes	% H.s	T1_call	T2_call	T3_call	B1_call	B2_call	B3_call	M_call	E1_call	E2_call	E3_call	Mk1_call	Mk2_call	Mk3_call
1700012H17Rik	RIKEN cDNA 1700012H17 gene	1	1/2	1	97	A	A	A	A	A	A	A	P	A	A	P	P	P	P
Gp9	glycoprotein 9 (platelet)	2		2	77	A	A	A	A	A	A	A	P	A	P	P	P	P	P
Pard3b	Par-3 partitioning defective 3 homolog B	3	1/4	3	87	A	A	A	A	A	A	A	P	A	A	P	P	P	P
Rnfl180	Ring finger protein 180	4		4	81	A	A	A	A	A	A	A	P	A	A	P	P	P	P
Serpina3g	serine (or cysteine) proteinase inhibitor, cl	5	5	5	63	P	P	P	P	P	P	P	P	P	P	A	P	P	P
Esam1	endothelial cell-specific adhesion molecu	6		6	77	A	A	A	A	A	A	A	P	P	P	P	P	P	P
Inpp4b	inositol polyphosphate-4-phosphatase, typ	7		7	92	A	P	A	A	P	A	A	P	A	A	P	P	P	P
Stxbp5	syntaxin binding protein 5 (tomosyn)	8	1/3	8	98	A	A	A	P	P	P	P	P	M	P	P	P	P	P
Gpr125	G protein-coupled receptor 125	9		9	96	A	A	A	A	A	A	A	P	A	P	P	P	P	P
Itga2b	integrin alpha 2b	10		10	85	P	P	P	P	P	M	P	P	P	P	P	P	P	P
Dapp1	Dual adaptor for phosphotyrosine and 3-ph	11	1/3	11	96	P	P	A	P	P	P	P	P	P	P	P	P	P	P
Pdzk1ip1	PDZK1 interacting protein 1	12	1/2	12	87	A	P	A	A	P	A	P	P	P	P	P	P	P	P
1110019K23Rik	RIKEN cDNA 1110019K23 gene	13	1/3	13	94	A	A	A	A	P	A	P	A	P	A	A	P	P	P
Ank	Progressive ankylosis	14	1/2	14	?	A	A	A	A		A	A	A	P	A	A	P	P	P
---	---	15		15	?	A	A	A	A	A	A	A	P	M	P	P	P	P	P
Pdzk1ip1	PDZK1 interacting protein 1	16	2/2		87	A	P	A	A	A	A	P	P	P	P	P	P	P	P
Rab13	RAB13, member RAS oncogene family	17		16	97	P	A	M	P	M	P	P	P	P	P	P	P	P	P
Tbxas1	thromboxane A synthase 1, platelet	18		17	85	A	A	A	A	A		P	P	P	P	P	P	P	P
1110019K23Rik	RIKEN cDNA 1110019K23 gene	19	2/3		94	A	A	A	P	P	A	P	P	P	P	P	P	P	P
Svil	supervillin	20	1/2	18	88	P	P	P	P	P	P	P	P	P	P	P	P	P	P
Dusp3	dual specificity phosphatase 3 (vaccinia virus	21		19	66	A	A	A	P	P	P	P	P	P	P	P	P	P	P
Angpt1	angiotensin I	22	1/3	20	99	P	P	A	P	P	P	P	P	P	A	P	P	P	P
Tek	endothelial-specific receptor tyrosine kina	23		21	95	P	A	A	A	A	A	P	P	A	P	P	P	P	P
Sri	Sorcin	24		22	?	A	A	A	A	A	A	A	A	A	A	M	A	M	M
Cxcl4	chemokine (C-X-C motif) ligand 4	25		23	74	P	P	A	A	P	P	P	P	P	P	P	P	P	P
Acadn	acetyl-Coenzyme A dehydrogenase, medu	26		24	91	P	P	P	P	P	P	P	P	P	P	P	P	P	P
Tuba8	tubulin, alpha 8	27	1/2	25	99	A	A	A	A	A	A	P	A	P	A	P	P	P	P
2610204M08Rik	RIKEN cDNA 2610204M08 gene	28	1/2	26	85	P	A	A	A	A	A	P	P	P	P	P	P	P	P
Cryz1	Crystallin, zeta (quinone reductase)-like 1	29	1/3	27	93	A	A	A	P	A	M	P	P	P	P	A	P	P	P
2610204M08Rik	RIKEN cDNA 2610204M08 gene	30	2/2		85	P	A	P	M	P	P	P	P	P	P	P	P	P	P
sept-11	septin 11	31	1/2	28	##	P	P	P	P	P	P	P	P	P	P	P	P	P	P
Ctla2b	cytotoxic T lymphocyte-associated protein	32		29	?	A	P	P	A	P	P	P	P	P	P	P	P	P	P
Angpt1	angiotensin I	33	2/3		99	A	M	A	P	A	A	P	A	A	P	A	P	P	P
Cd151	CD151 antigen	34	1/4	30	97	P	P	P	P	P	P	P	P	P	P	P	P	P	P
Cd151	CD151 antigen	35	2/4		97	P	P	P	P	P	P	P	P	P	P	P	P	P	P
Rab27b	RAB27b, member RAS oncogene family	36	1/2	31	97	A	A	A	A	A	A	P	P	P	P	P	P	P	P
Wdrx1	WD repeat domain, X-linked 1	37	1/2	32	99	P	P	P	P	P	P	P	P	P	P	P	P	P	P
Cd151	CD151 antigen	38	3/4		97	P	P	P	P	P	P	P	P	P	P	P	P	P	P
Trp1	transient receptor potential cation channel	39	1/2	33	##	A	P	A	A	A	A	A	A	A	A	A	A	A	P
Camk1	calcium/calmodulin-dependent protein kin	40	1/2	34	99	A	A	A	A	A	A	P	P	A	A	P	P	P	P
Myc1	myc target 1	41	1/4	35	87	A	A	A	A	A	A	A	P	A	A	P	P	P	P
Apln	apelin	42		36	83	P	P	P	P	A	A	P	A	A	P	P	P	P	P
Slk	STE20-like kinase (yeast)	43	3?		?	P	P	P	P	P	P	P	P	A	M	P	P	P	P
Tln1	tafin 1	44	1/2	38	99	P	P	P	P	P	P	P	P	P	P	P	P	P	P
Igf2	insulin-like growth factor 2	45		39	89	A	A	A	A	A	A	P	A	A	A	A	A	A	A
2700046G09Rik	RIKEN cDNA 2700046G09 gene	46	40		?	A	P	A	P	P	P	P	P	P	A	A	P	P	P
Itga6	integrin alpha 6	47	1/2	41	96	P	P	P	P	P	P	P	P	P	P	P	P	P	P
Ppp1r13b	Protein phosphatase 1, regulatory (inhibit	48	1/2	42	95	P	P	A	A	A	A	A	M	A	A	A	A	A	P
Atf7ip	activating transcription factor 7 interacti	49	43		?	P	P	P	P	P	P	P	P	P	A	P	P	P	P
Ikkbg	inhibitor of kappaB kinase gamma	50	1/3	44	89	P	P	P	P	P	P	P	P	P	P	P	P	P	P

Appendice D

Gènes spécifiques de M, Mk1 Mk2, Mk3

F1	log (fold change)	SAM d-score	P-value	FDR	T1	T2	T3	B1	B2	B3	M	E1	E2	E3	Mk1	Mk2	Mk3	Gene Symbol	
1435903 at	4,4	22,4	<0,001	0,01	-1,8	-1,3	-1,9	-1,5	-1,2	-1,3	3,4	-0,7	-0,7	-1,5	2,6	3,0	3,1	Cd300a	M
1416435 at	2,3	16,3	<0,001	0,01	0,0	-0,8	-0,7	-1,0	-1,1	-0,6	1,7	-0,7	-0,8	-1,0	1,7	1,8	1,5	Ltbr	M
1434705 at	4,1	15,3	<0,001	0,01	-1,1	-1,0	-1,6	-0,9	-1,1	-1,3	3,2	-1,2	-1,4	-1,9	3,3	2,7	2,4	Ctbp2	
1441811 x at	2,1	14,7	<0,001	0,01	0,0	-0,9	-0,7	-1,0	-0,7	0,0	1,5	-0,8	-0,8	-0,6	1,4	1,4	1,3	Tmem176a	
1440179 x at	4,2	14,7	<0,001	0,01	-1,2	-1,5	-1,5	-1,3	-1,1	-1,5	3,2	-1,1	-1,1	-0,9	3,4	2,6	2,2	Ibrdc1	
1430534 at	3,8	14,1	<0,001	0,01	-1,7	-1,3	-1,7	0,5	-0,9	-1,2	2,5	-1,0	-1,5	-1,4	2,9	2,2	2,5	Rnase6	
1450734 at	3,3	14,1	<0,001	0,01	-1,2	-0,9	-0,6	-1,4	-0,9	-0,8	2,5	-1,2	-0,9	-1,1	2,5	2,4	1,5	Lztr2	
1449164 at	4,7	14,0	<0,001	0,01	-1,3	-1,5	-1,6	-2,4	-2,5	-2,3	3,4	0,1	-0,2	-1,3	3,2	3,5	3,0	Cd68	M
1420980 at	3,6	12,6	<0,001	0,01	-0,7	-1,5	-1,3	-1,4	-1,1	-0,3	2,8	-0,9	-1,2	-1,7	2,9	2,7	1,9	Pak1	
1418972 at	0,7	12,2	<0,001	0,01	-0,2	-0,3	-0,4	0,1	-0,1	-0,2	0,6	-0,2	-0,1	-0,2	0,4	0,4	0,4	Bcl10	
1448727 at	2,3	11,9	<0,001	0,01	-0,7	-0,9	-0,3	0,3	-0,6	-1,3	1,6	-1,3	-0,6	-1,0	1,4	1,8	1,6	Tle6	
1455089 at	3,9	11,4	<0,001	0,01	-1,6	-1,8	-1,8	-0,7	0,1	-0,5	3,4	-1,6	-1,2	-2,2	3,0	2,4	2,4	Gng12	
1449222 at	4,0	11,2	<0,001	0,01	-1,4	-1,3	-1,2	-1,4	-1,2	-1,3	3,6	-0,7	-1,1	-1,6	2,2	3,1	2,2	Ebi3	M
1417070 at	2,3	11,2	<0,001	0,01	-0,9	-0,7	-0,7	-0,4	-0,5	-0,2	1,6	-0,9	-0,9	-0,9	1,3	1,9	1,2	Cyp4v3	
1422887 a at	3,6	10,9	<0,001	0,01	-1,4	-1,0	-1,0	-1,2	-1,4	-1,1	3,1	-1,0	-0,5	-1,2	2,7	2,3	1,7	Ctbp2	
1441997 at	1,4	10,7	<0,001	0,01	-0,2	-0,2	-0,7	-0,5	-0,7	-0,4	0,7	-0,6	-0,3	-0,3	1,1	0,9	1,2	Zfp184	
1440801 s at	4,4	10,6	<0,001	0,01	-1,4	-1,2	-1,4	-1,5	-1,6	-1,9	4,2	-0,7	-1,3	-1,1	2,3	2,9	2,6	Adrbk2	
1452514 a at	5,5	10,5	<0,001	0,01	-2,2	-2,1	-2,2	-2,2	-2,3	1,2	2,8	-1,7	-1,6	-2,2	3,8	4,6	4,0	Kit	M/Mk
1450070 s at	4,4	10,5	<0,001	0,01	-0,2	-1,7	-1,7	-1,8	-2,1	-0,2	3,6	-1,2	-1,6	-1,8	2,8	3,9	2,1	Pak1	
1425215 at	3,9	10,3	<0,001	0,01	-0,3	-1,5	-1,3	-1,6	-1,5	-1,1	1,7	-1,1	-1,2	-1,4	2,8	3,4	2,9	Ffar2	M
1437247 at	2,2	10,3	<0,001	0,01	-0,8	-0,5	-0,8	-0,7	-0,7	-0,8	1,8	-0,5	-0,5	-0,6	1,6	1,6	0,8	Fosl2	
1458438 at	3,0	10,2	<0,001	0,01	-1,6	-1,4	-1,6	-0,8	-1,0	0,8	2,4	-0,8	-1,3	-1,1	1,6	2,2	2,3	Ccdc122	
1415900 a at	4,0	10,0	<0,001	0,01	-1,5	-1,5	-1,3	-1,6	-1,5	0,0	2,1	-1,0	-1,1	-1,4	2,7	3,8	2,3	Kit	
1417071 s at	2,9	10,0	<0,001	0,01	-1,5	-1,5	-0,6	0,0	-0,4	0,3	2,0	-1,4	-1,6	-1,4	1,8	2,6	1,8	Cyp4v3	
1448355 at	2,7	10,0	<0,001	0,01	-0,9	-1,1	-0,2	-0,2	-0,7	-0,8	2,4	-1,3	-1,2	-1,2	2,0	1,8	1,3	Prss16	
1425603 at	3,8	10,0	<0,001	0,01	0,5	-1,8	-1,6	-1,9	-1,9	0,6	2,2	-1,5	-1,3	-2,0	3,0	3,0	2,6	Tmem176a	
1417936 at	6,4	10,0	<0,001	0,01	-1,7	-1,5	-1,9	-2,4	-2,0	-2,5	5,5	-1,5	-1,7	-2,3	4,5	4,9	2,6	Ccl9	M?
1419457 at	0,4	10,0	<0,001	0,01	-0,2	0,0	-0,2	-0,3	0,0	-0,3	0,3	0,1	-0,1	-0,2	0,4	0,3	0,4	Rgnef	
1423414 at	3,5	9,9	<0,001	0,01	-2,4	-1,6	-1,6	-0,5	-0,5	0,8	2,9	-0,2	-0,8	-1,6	2,4	1,8	2,8	Ptgs1	Mk
1416871 at	1,5	9,9	<0,001	0,01	-0,1	-0,3	-0,6	-0,5	-0,8	-0,4	0,7	-0,3	-0,8	-0,5	1,2	1,3	1,0	Adam8	Cd156
1423268 at	0,8	9,8	<0,001	0,01	-0,2	-0,2	-0,2	-0,4	-0,2	-0,3	0,8	-0,2	-0,2	-0,1	0,6	0,5	0,2	Itga5	Cd49e
1424265 at	2,5	9,8	<0,001	0,01	-1,1	-1,5	-1,5	-0,6	0,2	0,2	1,7	-1,3	-0,3	-1,1	1,8	2,0	1,4	Npl	
1419595 a at	3,2	9,7	<0,001	0,01	-1,3	-0,7	-0,9	-1,2	-0,2	-1,0	2,9	-1,2	-1,0	-1,3	2,2	2,2	1,5	Ggh	
1425739 at	0,8	9,6	<0,001	0,01	-0,1	-0,4	-0,1	-0,1	-0,4	-0,5	0,8	-0,3	-0,1	-0,3	0,5	0,5	0,3	Pld1	M
1449429 at	2,3	9,4	<0,001	0,01	-1,1	-0,5	-1,1	-1,0	-0,9	-0,6	1,4	-0,5	0,2	-0,5	2,1	1,2	1,3	Fkbp1b	
1436448 a at	3,7	9,2	<0,001	0,01	-1,4	-1,6	-1,5	-0,5	-0,8	-1,0	3,3	-0,8	-1,0	-1,4	2,1	1,6	2,9	Ptgs1	Mk
1454269 s at	6,5	9,1	<0,001	0,01	-2,6	-1,4	-2,5	-2,6	-2,6	-2,0	2,5	-2,5	-2,1	-0,9	5,3	6,3	5,1	P30519L02Rjk	4'
1418540 a at	3,0	9,1	<0,001	0,01	-0,6	-0,9	-1,0	-2,2	-2,7	-1,0	2,3	0,4	-0,5	0,0	2,3	2,0	2,0	Ptpre	
1418004 a at	4,5	9,1	<0,001	0,01	1,0	-1,7	-1,6	-2,3	-2,6	0,8	2,6	-2,2	-1,6	-2,5	3,6	3,5	3,0	Tmem176b	
1451690 a at	0,8	9,0	<0,001	0,01	-0,5	-0,3	-0,3	0,0	-0,3	-0,2	0,7	-0,1	-0,2	-0,2	0,5	0,6	0,4	Pvrl4	
1423909 at	3,3	9,0	<0,001	0,01	0,4	-1,3	-0,8	-1,1	-2,1	0,3	1,9	-1,5	-1,1	-1,6	2,5	2,6	1,9	Tmem176a	
1442916 at	2,5	9,0	<0,001	0,01	-1,0	-0,7	-0,9	-0,5	-0,5	-1,0	2,4	-0,7	-0,8	-0,8	1,1	1,7	1,7	Psd3	
1455019 x at	2,7	8,9	<0,001	0,01	-1,0	-1,2	-0,6	-0,7	-0,9	-0,8	2,6	-0,7	-0,8	-0,9	1,4	2,1	1,5	Ckap4	
1424542 at	5,2	8,9	<0,001	0,01	-1,7	-1,1	-0,4	-2,2	-2,1	-2,4	3,9	-1,0	-1,8	-1,7	4,0	4,5	2,1	S100a4	
1434518 at	2,2	8,9	<0,001	0,01	-1,8	-1,3	-1,1	0,0	-0,6	-0,7	1,4	0,0	-0,4	-0,2	1,5	2,0	1,4	Phka2	
1449215 at	0,7	8,8	<0,001	0,01	-0,2	-0,2	-0,3	-0,4	-0,3	0,1	0,8	-0,1	-0,5	0,0	0,4	0,3	0,5	Slc22a21	
1444408 at	1,4	8,7	<0,001	0,01	-0,7	-0,6	-0,9	0,0	-0,5	0,1	1,3	-0,3	-0,7	-0,1	0,8	1,1	0,6	30040O20Rjk	B
1454896 at	0,8	8,6	<0,001	0,01	-0,3	-0,4	-0,2	-0,1	-0,3	-0,5	0,6	0,2	-0,5	-0,1	0,5	0,5	0,4	Rbpsuh	
1451767 at	5,0	8,5	<0,001	0,01	-1,8	-2,2	-2,1	-0,6	-0,9	-1,7	3,6	-0,2	-2,0	-2,5	2,7	5,0	2,7	Ncf1	

Appendice D suite

Gene Symbol	Gene Title	ordre des probesets	probesets redondant	ordre des gènes	% de homologie (Hs)	T1_call	T2_call	T3_call	B1_call	B2_call	B3_call	M_call	E1_call	E2_call	E3_call	Mk1_call	Mk2_call	Mk3_call
Cd300a	CD300A antigen	1		1	53	A	P	A	A	A	A	P	A	M	A	P	P	P
Ltbr	lymphotoxin B receptor	2		2	73	P	A	A	A	A	A	P	A	A	A	P	P	P
Ctbp2	C-terminal binding protein 2	3	1/2	3	91	P	P	P	P	P	P	P	P	P	P	P	P	P
Tmem176a	transmembrane protein 176A	4	1/3	4	100	P	A	A	A	A	A	P	A	A	A	P	P	P
Ibrdc1	IBR domain containing 1	5		5	98	A	A	A	A	A	A	P	A	A	A	P	P	P
Rnase6	ribonuclease A family, member 6	6		6	100	A	P	A	P	P	A	P	P	A	A	P	P	P
Lztr2	leucine zipper transcription regulator 2	7		7	79	A	A	P	A	A	A	P	A	P	A	P	P	P
Cd68	CD68 antigen	8		8	75	P	P	P	P	P	P	P	P	P	P	P	P	P
Pak1	P21 (CDKN1A)-activated kinase 1	9	1/3	9	99	P	A	A	A	A	P	P	A	A	A	P	P	P
Bcl10	B-cell leukemia/lymphoma 10	10	1/2	10	95	P	P	P	P	P	P	P	P	P	P	P	P	P
Tle6	transducin-like enhancer of split 6	11		11	60	A	A	P	P	A	A	P	A	A	A	P	P	P
Gng12	guanine nucleotide binding protein	12		12	98	M	M	A	P	P	P	P	A	A	A	P	P	P
Ebi3	Epstein-Barr virus induced gene 3	13		13	66	A	A	A	A	A	A	P	A	A	A	P	P	P
Cyp4v3	cytochrome P450, family 4, subfamily v	14	1/2	14	77	A	A	A	M	P	P	P	A	A	A	P	P	P
Ctbp2	C-terminal binding protein 2	15	2/2	15	91	A	A	P	A	A	M	P	P	P	A	P	P	P
Zfp184	zinc finger protein 184 (Krueppel-like)	16		15	86	P	A	A	P	A	A	P	A	A	A	P	P	P
Adrbk2	adrenergic receptor kinase, beta 2	17		16	93	A	A	M	A	A	A	P	P	A	A	P	P	P
Kit	kit oncogene	18	1/3	17	90	M	A	A	A	A	P	P	A	M	A	P	P	P
Pak1	P21 (CDKN1A)-activated kinase 1	19	2/3	19	99	P	A	A	A	A	P	P	A	A	A	P	P	P
Ffar2	G protein-coupled receptor 43	20		18	89	P	A	A	A	A	A	P	A	A	A	P	P	P
Fosl2	fos-like antigen 2	21		19	98	A	A	A	A	A	A	P	A	A	A	P	P	P
Ccdc122	coiled-coil domain containing 122	22		20	72	A	A	A	A	A	P	P	P	A	A	P	P	P
Kit	kit oncogene	23	2/3	23	90	A	A	A	A	A	P	P	M	A	A	P	P	P
Cyp4v3	cytochrome P450, family 4	24	2/2	24	77	A	A	P	P	P	P	P	A	A	A	P	P	P
Prss16	protease, serine, 16 (thymus)	25		21	81	A	P	P	P	A	P	P	A	A	A	P	P	P
Tmem176a	transmembrane protein 176A	26	2/3	26	100	P	P	P	A	A	P	P	A	P	A	P	P	P
Ccl9	chemokine (C-C motif) ligand 9	27		23	?	P	P	M	M	P	A	P	P	P	A	P	P	P
Rgnef	Rho-guanine nucleotide exchange factor	28		24	86	A	A	A	A	A	A	A	A	A	A	A	A	A
Ptgs1	prostaglandin-endoperoxide synthase 1	29	1/2	25	93	A	A	A	P	A	A	P	A	A	A	P	P	P
Adam8	a disintegrin and metalloprotease domain 8	30		26	68	P	P	A	A	A	A	P	P	A	A	P	P	P
Itga5	integrin alpha 5	31	1/3	27	93	A	A	A	A	A	A	P	A	A	A	P	P	A
Npl	N-acetylneuraminatase pyruvate lyase	32		28	90	M	A	A	P	P	P	P	A	P	A	P	P	P
Ggh	gamma-glutamyl hydrolase	33		29	77	A	A	A	P	A	P	A	P	A	A	P	P	P
Pld1	phospholipase D1	34	1/3	30	?	A	A	P	M	A	P	A	P	A	A	P	P	A
Fkbp1b	FK506 binding protein 1b	35		31	98	A	A	A	A	A	A	P	A	A	A	P	P	P
Ptgs1	prostaglandin-endoperoxide synthase 1	36	2/2	36	93	A	P	A	P	P	P	P	P	A	A	P	P	P
930519L02R	RJKN cDNA 4930519L02 gene	37		32	?	A	P	A	A	A	P	P	A	P	P	P	P	P
Ptpre	protein tyrosine phosphatase,	38	1/3	33	96	P	P	P	P	P	P	P	P	P	P	P	P	P
Tmem176b	transmembrane protein 176B	39		34	62	P	P	P	P	A	P	P	P	P	A	P	P	P
Pvr14	Poliovirus receptor-related 4	40	1/2	35	93	A	A	A	A	A	A	M	A	A	A	A	A	A
Tmem176a	transmembrane protein 176A	41	3/3	41	100	P	A	A	A	A	P	P	A	A	A	P	P	P
Psd3	Pleckstrin and Sec7 domain containing 3	42		36	87	A	A	A	A	P	A	P	A	A	A	P	P	P
Ckap4	cytoskeleton-associated protein 4	43	1/4	37	88	A	A	A	M	P	A	M	P	A	A	A	P	P
S100a4	S100 calcium binding protein A4	44		38	95	P	P	P	A	A	P	P	P	A	P	P	P	P
Phka2	phosphorylase kinase alpha 2	45	1/5	39	95	P	P	P	P	P	P	P	P	P	P	P	P	P
Slc22a21	solute carrier family 22 member 21	46		40	81	A	A	A	A	A	A	P	A	A	M	A	P	P
130040O20R	RJKN cDNA B130040O20 gene	47		41	?	A	A	A	P	P	P	P	P	A	P	P	P	P
Rbpsuh	Recombining binding protein suppressor of H	48	1/2	42	99	P	P	P	P	P	P	P	P	P	P	P	P	P
Ncf1	neutrophil cytosolic factor 1	49	1/3	43	86	P	A	A	P	P	A	P	P	A	A	P	P	P

Appendice E

Gènes spécifiques de M, Mk1 Mk2, Mk3, B1, B2, B3

F1	log (fold change)	SAM d-score	P-value	FDR	T1	T2	T3	B1	B2	B3	M	E1	E2	E3	Mk1	Mk2	Mk3	Gene Symbol			
I448561_at	4,2	18,5	<0,001	<0,01	-2,6	-2,4	-2,1	2,0	1,6	1,6	3,0	-2,1	-2,2	-2,4	1,6	2,5	1,6	Ncf2	M	MK	
I418262_at	4,3	16,0	<0,001	<0,01	-2,4	-2,5	-2,8	2,8	2,4	2,6	1,5	-1,6	-2,2	-2,8	1,4	1,6	1,6	Syk	M	MK	B
I448213_at	5,4	14,1	<0,001	<0,01	-3,7	-2,4	-3,2	1,8	2,4	3,0	3,9	-2,4	-2,9	-3,4	2,5	2,8	1,6	Anxa1	M		
I418261_at	4,5	13,0	<0,001	<0,01	-2,4	-2,8	-2,8	2,9	2,5	2,8	1,4	-1,5	-2,0	-3,1	1,5	1,6	1,6	Syk	M	MK	B
I439389_s_at	4,0	12,8	<0,001	<0,01	-2,4	-2,7	-2,3	2,2	1,9	1,0	1,5	-1,2	-1,9	-2,4	2,4	2,6	1,6	Myadm	M		
I436100_at	4,9	12,2	<0,001	<0,01	-3,0	-2,4	-2,6	2,0	0,1	2,3	3,4	-2,5	-2,7	-2,4	2,8	2,5	1,6	Sh2d5			
I425797_a_at	3,6	12,2	<0,001	<0,01	-2,0	-2,1	-2,0	2,4	2,3	2,6	1,2	-2,1	-1,8	-2,0	0,7	1,3	1,6	Syk	M	MK	B
I429114_at	1,9	11,1	<0,001	<0,01	-1,2	-1,2	-1,1	0,3	0,7	1,0	1,6	-0,9	-1,1	-0,9	1,2	1,0	1,6	Sestd1			
I423321_at	2,8	11,1	<0,001	<0,01	-1,5	-1,8	-1,8	1,7	1,6	0,8	0,9	-1,1	-1,4	-1,4	2,3	1,4	1,6	Myadm	M		
I430373_at	3,1	10,5	<0,001	<0,01	-1,7	-1,5	-1,9	1,3	1,1	1,3	3,0	-1,2	-1,7	-1,8	0,9	0,8	1,6	5430427019Rik			
I420620_a_at	1,4	10,2	<0,001	<0,01	-0,7	-0,3	-0,6	0,5	0,6	0,9	0,9	-0,7	-1,1	-1,1	0,5	0,6	1,6	Rnfl3			
I452410_a_at	3,5	10,2	<0,001	<0,01	-2,3	-1,7	-1,7	1,8	0,5	1,0	3,0	-1,4	-2,0	-1,9	1,4	2,2	1,6	Fes	M	MK	B
I448611_at	1,8	10,2	<0,001	<0,01	-1,2	-1,3	-1,0	0,6	0,6	1,2	1,0	-0,9	-0,8	-0,9	1,1	1,1	1,6	Wwc2			
I419493_a_at	3,6	10,1	<0,001	<0,01	-2,1	-2,0	-2,4	1,7	2,1	1,0	3,0	-1,5	-2,1	-1,9	0,1	2,3	1,6	Tpd52			
I419515_at	2,9	10,1	<0,001	<0,01	-1,8	-2,3	-0,9	1,6	1,8	0,7	2,1	-1,3	-1,5	-1,9	0,6	1,2	1,6	Fgd2			
I418747_at	3,3	9,9	<0,001	<0,01	-2,6	-1,7	-1,7	1,4	0,9	1,1	3,0	-1,7	-1,4	-1,8	1,2	2,1	1,6	Sfpil(Pu.1)	M	MK	B
I435582_at	5,4	9,6	<0,001	<0,01	-3,7	-3,9	-3,2	4,0	3,7	3,1	1,3	-1,4	-2,0	-3,3	1,7	1,9	1,6	Wdfy4			
I427368_x_at	3,7	9,5	<0,001	<0,01	-2,7	-2,5	-1,8	1,9	0,6	1,2	3,1	-1,0	-2,0	-2,1	1,9	2,1	1,6	Fes	M	MK	B
I428416_at	2,0	9,4	<0,001	<0,01	-1,1	-0,9	-0,9	0,8	1,3	1,1	-0,2	-1,1	-0,9	-1,4	1,2	1,2	1,6	5110050N22Rik			
I460231_at	2,8	9,3	<0,001	<0,01	-1,4	-1,7	-1,5	1,7	1,8	1,4	2,5	-1,3	-1,4	-1,7	0,0	0,6	1,6	Irf5			B
I435477_s_at	5,0	9,2	<0,001	<0,01	-2,2	-3,2	-2,8	2,3	2,6	1,1	4,5	-2,2	-2,2	-3,7	0,8	2,1	1,6	Fcgr2b	M		B
I419494_a_at	2,7	9,1	<0,001	<0,01	-1,5	-1,6	-1,4	1,3	1,5	0,7	2,4	-1,3	-1,5	-1,5	-0,1	1,7	1,6	Tpd52			B
I457239_at	2,1	9,0	<0,001	<0,01	-1,2	-1,4	-1,5	1,9	1,4	1,2	0,4	-0,9	-0,9	-1,2	0,0	1,0	1,6	Syk	M	MK	B
I435141_at	1,0	8,9	<0,001	<0,01	-0,7	-0,7	-0,7	0,3	0,5	0,4	0,4	-0,1	-0,4	-0,7	0,7	0,6	1,6	Sft2d2			
I455549_at	2,0	8,7	<0,001	<0,01	-0,9	-1,1	-1,5	0,1	0,5	0,7	1,0	-1,0	-1,1	-1,0	1,8	1,3	1,6	Sestd1			
I427304_at	1,2	8,6	<0,001	<0,01	-0,6	-0,9	-0,6	0,8	0,3	0,3	0,7	-1,1	-0,4	-0,5	1,1	0,5	1,6	Vps18			
I418294_at	4,2	8,4	<0,001	<0,01	-2,4	-2,1	-2,5	3,0	3,6	2,0	-0,5	-2,2	-2,3	-2,6	2,6	2,3	1,6	Epb4.14b			
I416318_at	5,1	8,3	<0,001	<0,01	-4,5	-3,2	-3,4	1,3	1,6	2,6	4,1	-1,0	-1,6	-2,8	2,5	2,4	1,6	Serpina1a			
I448547_at	1,5	8,3	<0,001	<0,01	-0,5	-1,1	-0,7	0,8	0,8	1,1	1,1	-1,1	-1,0	-0,7	0,4	0,6	1,6	Rassf3			
I425538_x_at	3,9	8,2	<0,001	<0,01	-1,4	-2,3	-2,6	2,9	2,8	2,6	2,1	-1,7	-2,6	-2,1	0,2	1,5	1,6	Ceacam1	M	MK	B
I428156_at	3,1	8,2	<0,001	<0,01	-1,8	-2,0	-0,1	1,7	1,8	1,7	2,3	-2,1	-2,0	-2,3	1,6	0,3	1,6	Gng2			
I434245_a_at	2,5	8,1	<0,001	<0,01	-1,6	-1,8	-1,9	1,4	1,9	2,1	0,8	-0,6	-1,5	-1,0	0,3	1,0	1,6	Cybasc3			
I450276_a_at	5,9	8,1	<0,001	<0,01	-3,3	-2,9	-3,6	4,6	2,6	3,7	2,7	-3,3	-2,6	-3,9	4,4	-1,1	1,6	Scin		MK	
I456046_at	5,7	8,1	<0,001	<0,01	-2,2	-3,7	-4,2	4,1	3,9	4,1	3,7	-2,3	-2,2	-4,1	0,9	0,6	1,6	C1qrl(cd93)	M		
I451941_a_at	4,2	8,0	<0,001	<0,01	-2,1	-2,4	-2,3	1,9	2,5	0,8	4,4	-2,1	-2,2	-2,6	-0,2	1,8	1,6	Fcgr2b	M		B
I426926_at	4,7	7,5	<0,001	0,01	-3,4	-3,4	-2,4	2,9	2,0	2,2	1,5	0,2	-2,9	-3,3	1,9	3,0	1,6	Plec2			
I421375_a_at	4,2	7,5	<0,001	0,01	-3,0	-2,0	-1,2	0,6	1,2	1,9	3,7	-2,1	-2,7	-3,1	3,4	2,9	1,6	SI00a6			
I423891_at	1,7	7,4	<0,001	0,01	-0,5	-1,2	-0,4	1,0	0,8	0,3	0,8	-1,2	-1,3	-1,0	0,8	1,4	1,6	Gstt3			
I435476_a_at	2,5	7,4	<0,001	0,01	-1,5	-1,4	-1,6	1,3	1,4	0,4	2,8	-1,1	-1,4	-1,3	0,0	0,9	1,6	Fcgr2b			
I419589_at	5,3	7,4	<0,001	0,01	-2,2	-3,7	-3,3	3,9	3,6	4,2	3,6	-2,8	-2,2	-3,0	0,5	0,4	1,6	C1qrl(cd93)		MK	
I427630_x_at	3,3	7,3	<0,001	0,01	-1,5	-2,5	-2,0	2,5	2,6	2,3	1,7	-1,0	-2,1	-1,8	0,0	1,4	1,6	Ceacam1	M	MK	B
I455332_x_at	3,5	7,3	<0,001	0,01	-2,0	-1,8	-2,0	1,8	2,1	0,8	4,0	-2,0	-1,7	-2,5	-0,1	1,3	1,6	Fcgr2b	M		B
I444143_at	1,2	7,3	<0,001	0,01	-0,4	-0,6	-0,5	0,4	0,3	1,3	0,2	-0,8	-0,7	-0,7	0,9	0,2	1,6	3230342M21Rik			F
I434839_s_at	0,5	7,3	<0,001	0,01	-0,5	-0,4	-0,4	0,4	0,2	0,2	0,2	-0,4	0,0	-0,3	0,5	0,1	1,6	Tbllx1			
I426120_a_at	2,5	7,2	<0,001	0,01	-1,7	-1,3	-1,2	0,4	2,3	2,4	0,5	-1,1	-1,4	-1,6	0,2	1,5	1,6	Cd244			
I423768_at	3,2	7,1	<0,001	0,01	-1,4	-1,8	-1,8	2,4	1,8	2,3	2,5	-1,4	-1,6	-1,9	-1,1	0,9	1,6	Unc93b1			
I454895_at	1,8	7,1	<0,001	0,01	-1,4	-1,4	-1,3	1,3	1,4	1,8	0,6	-0,6	-0,7	-0,7	0,2	0,4	1,6	Cybasc3			
I453474_at	2,2	7,1	<0,001	0,01	-0,4	-0,6	-1,7	1,5	1,9	1,0	1,6	-0,8	-1,8	-1,9	0,0	1,3	1,6	1300007F04Rik			
I429831_at	5,2	7,0	<0,001	0,01	-3,7	-3,7	-3,9	4,3	3,7	3,5	2,1	-1,0	-1,9	-2,7	0,0	1,8	1,6	Pik3ap1			

Appendice E suite

Gene Symbol	Gene Title	ordre des probesets	probesets redondants	ordre des gènes	% de homologie (Hs)	T1_call	T2_call	T3_call	B1_call	B2_call	B3_call	M_call	E1_call	E2_call	E3_call	Mk1_call	Mk2_call	Mk3_call
Ncf2	neutrophil cytosolic factor 2	1		1	86	A	P	P	P	P	P	P	P	P	A	P	P	P
Syk	spleen tyrosine kinase	2	1/4	2	94	P	A	A	P	P	P	P	P	A	A	P	P	P
Anxa1	annexin A1	3		3	91	P	P	P	P	P	P	P	P	P	P	P	P	P
Syk	spleen tyrosine kinase	4	2/4	4	94	A	A	A	P	P	P	P	P	A	A	P	P	P
Myadm	myeloid-associated differentiation mark	5	1/2	4	89	P	M	A	P	P	P	P	P	P	A	P	P	P
Sh2d5	SH2 domain containing 5	6	1/2	5	88	A	P	A	P	P	P	P	A	A	A	P	P	P
Syk	spleen tyrosine kinase	7	3/4	7	94	P	A	A	P	P	P	P	A	M	A	P	P	P
Sestd1	SEC14 and spectrin domains 1	8	1/2	6	99	A	P	P	P	P	P	P	P	A	P	P	P	P
Myadm	myeloid-associated differentiation mark	9	2/2	9	89	A	A	A	P	P	P	P	A	A	A	P	P	P
5430427019Ri	RIKEN cDNA 5430427019 gene	10		7	83	A	A	A	P	P	P	P	A	A	A	P	P	P
Rnf13	ring finger protein 13	11		8	97	P	P	P	P	P	P	P	P	P	P	P	P	P
Fes	feline sarcoma oncogene	12	1/2	9	93	A	A	P	P	P	P	P	P	A	A	P	P	P
Wwc2	WW C2 and coiled-coil domain cont 2	13		10	91	A	A	A	P	P	P	P	A	A	A	P	P	P
Tpd52	tumor protein D52	14	1/3	11	92	P	P	P	P	P	P	P	P	P	P	P	P	P
Fgd2	FYVE, RhoGEF and PH domain containi	15		12	87	A	A	A	P	P	P	P	A	A	A	P	P	P
Sfp1l(Pu.1)	SFFV proviral integration 1	16		13	95	A	A	P	P	P	P	P	P	P	P	P	P	P
Wdfy4	WD repeat and FYVE domain containi	17		14	?	A	A	A	P	P	P	P	P	M	A	P	P	P
Fes	feline sarcoma oncogene	18	2/2	93	A	A	A	P	P	P	P	M	A	A	A	P	P	P
3110050N22Ri	RIKEN cDNA 3110050N22 gene	19	1/3	15	93	A	A	A	P	P	P	M	A	A	A	P	P	P
Irf5	interferon regulatory factor 5	20		16	94	P	P	P	P	P	P	P	P	A	A	P	P	P
Fcgr2b	Fc receptor, IgG, low affinity IIb	21	1/4	17	62	P	P	P	P	P	P	P	P	P	A	A	P	P
Tpd52	tumor protein D52	22	2/3	92	A	A	A	P	P	P	P	P	A	A	A	P	P	P
Syk	Spleen tyrosine kinase	23	4/4	94	A	A	A	P	P	P	P	P	A	A	A	P	P	P
Sft2d2	SFT2 domain containing 2	24	1/2	18	92	P	P	P	P	P	P	P	P	P	P	P	P	P
Sestd1	SEC14 and spectrin domains 1	25	2/2	99	P	A	P	P	P	P	P	P	P	A	A	P	P	P
Vps18	vacuolar protein sorting 18 (yeast)	26	1/2	19	97	M	A	A	P	P	P	P	A	P	P	P	P	P
Epb4.1l4b	erythrocyte protein band 4.1-like 4b	27		20	94	P	P	A	P	P	P	P	P	A	P	P	P	P
Serpinb1a	serine (or cysteine) peptidase inhibitor,	28		21	84	A	A	A	P	P	P	P	M	P	A	P	P	P
Rassf3	Ras association (RalGDS/AF-6) domain	29	1/3	22	95	P	P	P	P	P	P	P	P	P	P	P	P	P
Ceacam1	CEA-related cell adhesion molecule 1	30	1/7	23	64	P	A	A	P	P	P	P	P	A	A	P	P	P
Gng2	guanine nucleotide binding protein (G	31	1/4	24	100	A	P	P	P	P	P	P	M	A	A	P	P	P
Cybasc3	cytochrome b, ascorbate dependent 3	32	1/2	25	86	P	P	P	P	P	P	P	P	P	P	P	P	P
Scin	scinderin	33		26	64	P	P	A	P	P	P	P	P	P	A	P	P	P
C1qr1(cd93)	complement component 1, q subcompo	34	1/4	27	75	P	A	P	P	P	P	P	P	P	A	P	P	P
Fcgr2b	Fc receptor, IgG, low affinity IIb	35	2/4	62	P	A	P	P	P	P	P	P	P	A	A	P	P	P
Plcg2	phospholipase C, gamma 2	36		28	96	P	A	P	P	P	P	P	P	P	P	P	P	P
S100a6	S100 calcium binding protein A6 (calcy	37		29	99	A	P	P	P	P	P	P	P	P	A	P	P	P
Gstt3	glutathione S-transferase, theta 3	38		30	80	P	A	P	P	P	P	P	A	A	A	P	P	P
Fcgr2b	Fc receptor, IgG, low affinity IIb	39	3/4	62	P	P	A	P	P	P	P	M	M	P	P	P	P	P
C1qr1(cd93)	complement component 1, q subcompo	40	2/4	75	P	A	A	P	P	P	P	P	P	A	A	P	P	P
Ceacam1	CEA-related cell adhesion molecule 1	41	2/7	31	64	P	A	A	P	P	P	P	P	A	A	P	P	P
Fcgr2b	Fc receptor, IgG, low affinity IIb	42	4/4	62	P	A	A	P	P	P	P	P	A	A	A	P	P	P
3230342M21Ri	RIKEN cDNA B230342M21 gene	43		32	92	A	A	A	P	A	P	M	A	A	A	A	M	A
Tb1l1xr1	transducin (beta)-like 1X-linked recepto	44		33	99	P	P	P	P	P	P	P	P	P	P	P	P	P
Cd244	CD244 natural killer cell receptor 2B4	45	1/2	34	54	A	A	A	P	P	P	P	A	A	A	P	P	P
Unc93b1	unc-93 homolog B1 (C. elegans)	46		35	94	M	A	A	P	P	P	P	P	A	A	A	P	P
Cybasc3	cytochrome b, ascorbate dependent 3	47	2/2	86	P	P	P	P	P	P	P	P	P	P	P	P	P	P
1300007F04Ri	RIKEN cDNA 1300007F04 gene	48		36	90	P	P	A	P	P	P	P	P	A	A	P	P	P
Pik3ap1	phosphoinositide-3-kinase adaptor prot	49	1/2	37	89	P	A	A	P	P	P	P	P	A	A	P	P	P

Appendice F

[illegible]

Appendice G

Gènes spécifiques de E1, E2, E3																																									
F1	log (fold change)	SAM d-score	P-value	FDR	T1	T2	T3	B1	B2	B3	M	E1	E1-C	E2	E2-C	E3	E3-C	Mk1	Mk2	Mk3	Gene Symbol																				
I454265_a_at	2.8	39.3	<0.001	0.01	-0.9	-0.7	-0.5	-0.6	-0.8	-0.8	-0.6	2.1	-0.8	2.1	-0.7	2.3	-0.6	-0.3	-0.8	-0.5	Abcb10	E																			
I450512_at	2.2	33.4	<0.001	0.01	-0.6	-0.6	-0.4	-0.4	-0.6	-0.6	-0.5	1.7	0.9	1.7	1.0	1.7	1.0	-0.5	-0.8	-0.1	Ntn4																				
I427221_at	5.3	32.8	<0.001	0.01	-1.4	-1.3	-1.3	-1.2	-1.2	-1.1	-1.4	4.2	0.4	3.7	-0.1	4.2	0.4	-1.2	-1.5	-0.4	Xtrp3s1																				
I417337_at	5.1	29.4	<0.001	0.01	-1.4	-1.1	-1.2	-1.3	-1.4	-1.6	-1.4	3.8	-2.4	4.0	-2.2	4.0	-2.2	-1.2	-1.2	0.2	Epb4_2	E																			
I435435_at	2.9	29.3	<0.001	0.01	-0.8	-0.8	-0.4	-0.6	-0.8	-0.7	-0.8	2.0	2.9	2.3	3.2	2.2	3.1	-0.3	-0.7	-0.6	Cttnbp2																				
I419069_at	2.9	25.0	<0.001	0.01	-0.5	-0.8	-0.4	-0.9	-1.3	-1.2	-0.8	2.2	1.0	2.3	1.1	2.3	1.1	-0.4	-0.4	-0.1	Rabgef1																				
I427357_at	6.0	25.0	<0.001	0.01	-1.5	-1.6	-1.6	-1.7	-1.8	-0.6	-1.3	4.3	5.6	4.8	6.1	4.9	6.3	-1.6	-1.7	-0.7	Cda																				
I422754_at	3.7	22.8	<0.001	0.01	-1.2	-1.3	-0.7	-1.4	-1.1	-1.4	-0.9	3.1	-0.8	3.0	-0.9	2.9	-1.0	-0.9	-0.2	0.2	Tmod1	E																			
I437959_at	2.9	22.4	<0.001	0.01	-0.6	-0.9	-0.3	-0.5	-0.5	-0.4	-0.8	1.9	3.1	2.3	3.5	2.0	3.2	-0.9	-0.7	-0.7	Zfp324																				
I422493_at	3.4	22.1	<0.001	0.01	-1.0	-1.2	-0.3	-0.5	-0.9	-1.0	-1.6	2.7	-0.7	2.7	-0.8	2.9	-0.6	-0.7	-1.2	-0.1	Cpox	E																			
I454622_at	6.5	21.7	<0.001	0.01	-1.7	-1.6	-2.2	-2.0	-1.7	-1.8	-1.7	4.6	-0.8	5.1	-0.3	5.3	-0.1	-1.2	-1.5	0.3	Slc38a5	E																			
I416741_at	4.0	21.1	<0.001	0.01	-0.9	-1.0	-1.0	-1.4	-1.1	-1.3	-1.0	2.9	2.0	3.4	2.5	2.9	2.0	-0.3	-1.2	0.0	Col5a1																				
I429168_at	2.3	19.4	<0.001	0.01	-0.8	-0.6	-0.6	-0.7	-0.7	-0.6	-0.6	1.6	0.1	1.8	0.3	1.8	0.3	-0.1	-0.4	0.0	Btdb4																				
I416075_at	1.2	19.3	<0.001	0.01	0.2	-0.5	-0.2	-0.2	-0.3	-0.1	-0.2	1.0	0.8	0.9	0.6	0.9	0.6	-0.3	-0.6	-0.4	Savi1																				
I417806_at	4.1	19.3	<0.001	0.01	-1.2	-1.2	-0.6	-0.9	-1.3	-0.8	-1.0	3.3	1.0	2.8	0.4	3.1	0.8	-0.6	-1.2	-0.4	Popdc2																				
I430632_at	2.8	19.1	<0.001	0.01	-0.8	-0.7	-0.6	-0.7	-0.8	-0.6	-0.8	2.2	-0.1	2.3	0.0	1.9	-0.4	-0.6	-0.9	0.1	Asb1																				
I416994_at	1.0	19.1	<0.001	0.01	-0.1	0.1	0.0	-0.4	-0.4	-0.4	-0.3	0.8	1.5	0.8	1.5	0.8	1.5	-0.4	-0.1	-0.3	Tcl1																				
I451395_at	3.2	18.9	<0.001	0.01	-0.4	-0.2	-1.0	-1.3	-1.2	-1.4	-0.9	2.6	3.0	2.6	3.0	2.6	3.0	-1.2	-0.4	0.3	Ccdc92																				
I424437_s_at	3.0	18.9	<0.001	0.01	-0.7	-0.5	0.0	-0.9	-0.7	-1.0	-1.1	2.5	-1.4	2.1	-1.8	2.2	-1.7	-0.5	-1.2	-0.3	Abcg4																				
I416402_at	3.6	18.8	<0.001	0.01	-1.3	-1.2	-1.2	-1.4	-1.6	-0.9	-0.1	2.8	-0.6	2.7	-0.7	2.8	-0.6	-0.3	-0.6	0.2	Abcb10	E																			
I448537_at	0.9	18.7	<0.001	0.01	-0.1	-0.1	0.0	-0.4	-0.3	-0.3	-0.4	0.8	1.2	0.8	1.2	0.7	1.2	-0.2	-0.2	0.0	Tcl1																				
I425227_a_at	3.5	18.3	<0.001	0.01	-1.4	-1.5	-1.5	-0.5	-1.1	-0.6	-0.1	2.8	0.1	2.6	0.0	2.6	-0.1	-0.5	-0.7	0.0	Atps60a1																				
I417638_at	2.3	18.2	<0.001	0.01	-0.6	-1.0	-1.1	-0.4	-0.4	-0.8	-0.6	1.8	1.3	1.7	1.2	1.8	1.4	-0.2	-0.4	0.3	Lefty1																				
I425677_a_at	4.3	18.1	<0.001	0.01	-1.5	-1.3	-1.3	-1.5	-1.1	-0.9	-1.0	3.1	-1.1	3.6	-0.5	3.3	-0.9	-0.5	-1.1	0.1	Ank1	E																			
I419068_at	2.6	18.1	<0.001	0.01	-1.2	-0.8	-0.5	-0.9	-0.7	-1.0	-0.7	1.9	1.1	2.1	1.3	2.2	1.4	-0.5	-0.3	0.2	Rabgef1																				
I422524_at	2.9	17.9	<0.001	0.01	-0.4	-0.8	-0.1	-0.7	-0.5	-0.7	-0.9	2.2	-0.6	2.5	-0.3	2.4	-0.4	-1.1	-1.5	-0.2	Abcb6	E																			
I451667_at	1.3	17.7	<0.001	0.01	-0.3	-0.7	0.0	-0.3	-0.3	-0.1	-0.4	1.1	0.7	1.0	0.6	1.1	0.7	-0.5	-0.4	0.0	S30043G21RikCS																				
I416403_at	3.9	17.7	<0.001	0.01	-1.3	-1.2	-1.1	-1.7	-1.9	-1.2	-0.3	3.1	-0.8	3.1	-0.8	2.9	-0.9	-0.1	-0.5	0.2	Abcb10	E																			
I419063_at	4.7	17.7	<0.001	0.01	-1.4	-1.2	-1.3	-1.1	-1.2	-1.1	-1.2	3.1	2.7	3.9	3.5	3.9	3.4	-0.7	-0.7	-0.9	Ugt8																				
I450556_at	4.1	17.6	<0.001	0.01	-1.4	-1.3	-0.8	-1.5	-1.5	-1.2	-0.1	3.4	-1.3	3.2	-1.5	3.2	-1.5	-0.7	-1.1	0.7	Spnb1	E																			
I451453_at	3.1	17.6	<0.001	0.01	-0.8	-0.8	-0.7	-1.0	-1.0	-0.9	-0.2	2.3	-1.3	2.4	-1.2	2.8	-0.8	-0.9	-0.9	-0.2	Dapk2	E																			
I439794_at	3.9	17.6	<0.001	0.01	-1.0	-0.9	-0.5	-1.1	-1.1	-1.1	-1.1	2.6	1.2	3.2	1.8	3.2	1.8	-1.0	-1.0	-0.3	---																				
I435389_at	3.0	17.5	<0.001	0.01	-0.6	-0.8	-0.9	-0.8	-0.9	-0.8	-0.5	2.0	0.8	2.3	1.1	2.6	1.3	-0.7	-0.2	-0.7	Reps2																				
I422492_at	3.8	17.4	<0.001	0.01	-1.6	-1.2	-0.2	-0.5	-0.8	-1.0	-1.5	2.7	-0.7	3.1	-0.2	3.2	-0.2	-1.0	-0.7	-0.4	Cpox	E																			
I417165_at	1.6	17.3	<0.001	0.01	-0.7	-0.1	-0.4	0.0	-0.1	-0.3	-0.6	1.3	0.9	1.3	0.9	1.3	0.9	-0.6	-0.7	-0.2	Mbd2	E																			
I417338_at	4.9	17.3	<0.001	0.01	-1.7	-1.5	-1.4	-1.5	-1.7	-1.7	-1.2	3.7	-2.2	3.8	-2.1	3.9	-2.0	-1.2	-1.0	1.4	Epb4_2	E																			
I418592_at	3.3	17.3	<0.001	0.01	-1.1	-1.3	-0.4	-1.2	-0.8	-0.8	-1.2	2.5	-0.5	2.6	-0.4	2.7	-0.3	-0.7	-0.1	-0.1	Dnaja4																				
I456891_at	3.7	17.3	<0.001	0.01	-0.8	-1.3	-1.0	-1.2	-1.2	-0.9	-0.9	3.2	1.0	2.6	0.3	2.9	0.7	-1.4	-0.7	0.6	Dennd2c																				
I417636_at	2.6	17.1	<0.001	0.01	-0.1	-1.0	-0.6	-0.6	-0.8	-1.2	-1.0	2.0	-2.3	2.3	-2.0	2.0	-2.3	-0.6	-0.6	0.1	Slc6a9																				
I438774_s_at	2.2	16.9	<0.001	0.01	-0.8	-0.1	-0.3	-0.4	-0.6	-0.8	-0.7	1.9	1.7	1.5	1.4	1.5	1.4	-0.3	-0.7	-0.2	Pgm21l																				
I460358_s_at	1.6	16.8	<0.001	0.01	-0.4	-0.4	-0.5	-0.7	-0.6	-0.8	0.3	1.2	1.6	1.4	1.8	1.4	1.8	-0.3	-0.4	-0.1	Nudt22																				
I449469_at	3.0	16.8	<0.001	0.01	-1.2	-0.6	-0.6	-0.6	-0.1	-1.2	-1.1	2.2	-0.9	2.6	-0.5	2.2	-0.9	-0.5	-0.6	-0.5	Pkd21l2																				
I422268_at	0.9	16.7	<0.001	0.01	-0.3	-0.5	-0.7	0.0	-0.1	-0.2	-0.1	0.8	0.6	0.8	0.6	0.7	0.5	-0.1	-0.3	0.0	Rps6kb2																				
I427140_at	2.0	16.6	<0.001	0.01	-0.6	-0.7	-0.5	-0.7	-0.1	-0.2	-0.9	1.7	2.6	1.5	2.3	1.8	2.6	-0.1	-0.7	-0.3	Pvt1																				
I429019_s_at	0.9	16.5	<0.001	0.01	-0.1	-0.4	-0.2	-0.4	-0.1	-0.3	-0.2	0.8	0.0	0.8	0.0	0.8	0.0	-0.7	-0.1	0.0	Pon2																				
I423580_at	2.5	16.5	<0.001	0.01	-0.7	-0.6	-0.5	-0.9	-0.6	-0.5	-0.4	2.6	1.7	2.1	2.3	1.8	1.9	-0.4	-0.9	-0.2	Ankrd48																				
I418591_at	4.8	16.4	<0.001	0.01	-2.1	-1.5	-1.2	-1.5	-1.4	-1.5	-2.0	3.5	-0.7	3.7	-0.5	3.9	-0.3	-0.1	-0.5	0.7	Dnaja4																				
I424135_at	1.4	16.3	<0.001	0.01	-0.4	-0.2	-0.1	-0.4	-0.5	-0.6	-0.6	1.2	2.1	1.2	2.0	1.0	1.9	0.0	-0.3	-0.4	Rspry1																				
I450425_a_at	1.1	16.1	<0.001	0.01	-0.5	-0.3	-0.2	-0.5	-0.5	-0.4	0.0	1.0	0.6	0.9	0.6	0.9	0.5	-0.3	-0.3	0.1	Copz1																				
I460866_at	1.4	16.1	<0.001	0.01	0.1	-0.7	-0.1	-0.7	-0.6	-0.4	-0.1	1.0	0.4	1.1	0.4	1.1	0.4	-0.4	-0.2	-0.3	70004717Rik																				

Appendice G suite

Gene Symbol	Gene Title	GI	ordre des probesets	probesets redondants	ordre des gènes	% de homologie (Hs)	T1_call	T2_call	T3_call	B1_call	B2_call	B3_call	M_call	E1_call	E2_call	E3_call	Mk1_call	Mk2_call	Mk3_call
Abcb10	ATP-binding cassette, sub-family B (MDR/TA	↑	1	1/3	1	85	A	A	A	A	A	A	A	P	P	P	P	P	A
Ntn4	netrin 4	-	2		2	?	A	A	A	A	A	A	A	P	P	P	P	A	A
Xtrp3s1	X transporter protein 3 similar 1 gene	-	3		3	91	A	A	A	A	A	A	A	P	P	P	P	A	A
Epb4.2	erythrocyte protein band 4.2	→	4	1/2	4	78	A	A	A	A	A	A	A	P	P	P	P	A	P
Cttnbp2	cortactin binding protein 2	-	5		5	88	A	A	A	A	A	A	A	P	P	P	P	A	P
Rabgef1	RAB guanine nucleotide exchange factor (GEF	-	6	1/3	6	95	P	P	P	P	P	P	P	P	P	P	P	P	P
Cda	Cytidine deaminase	-	7		7	85	A	A	A	A	M	A	P	P	P	P	A	A	M
Tmod1	tropomodulin 1	-	8	1/2	8	98	P	P	P	A	P	A	P	P	P	P	P	P	P
Zfp324	zinc finger protein 324	-	9		9	80	P	A	P	A	M	P	P	P	P	P	A	A	A
Cpx	coproporphyrinogen oxidase	-	10	1/2	10	88	P	P	P	P	P	P	P	P	P	P	P	P	P
Slc38a5	solute carrier family 38, member 5	↑	11		11	89	A	A	A	A	A	A	A	P	P	P	A	A	P
Col5a1	procollagen, type V, alpha 1	→	12	1/3	12	95	A	A	A	A	A	A	A	P	P	P	A	A	P
Btd4	BTB (POZ) domain containing 4	-	13	1/4	13	93	A	A	A	A	A	A	A	P	P	P	A	A	P
Sav1	salvador homolog 1 (Drosophila)	-	14	1/2	14	96	P	P	P	P	P	P	P	P	P	P	P	P	P
Popdc2	popeye domain containing 2	-	15		15	85	A	A	P	A	A	A	A	P	P	P	A	A	M
Asb1	ankyrin repeat and SOCS box-containing prote	-	16	1/3	16	95	A	A	A	A	A	A	P	P	P	P	A	A	A
Ttc1	tetratricopeptide repeat domain 1	-	17	1/2	17	92	P	P	P	P	P	P	P	P	P	P	P	P	P
Ccd92	coiled-coil domain containing 92L	-	18		18	91	P	P	P	P	P	A	P	P	P	P	A	P	P
Abcg4	ATP-binding cassette, sub-family G (WHITE)	-	19	1/2	19	98	A	A	P	A	A	A	A	P	P	P	M	A	A
Abcb10	ATP-binding cassette, sub-family B (MDR/TA	↑	20	2/3		85	P	P	P	P	P	P	P	P	P	P	P	P	P
Ttc1	tetratricopeptide repeat domain 1	-	21	2/2		92	P	P	P	P	P	P	P	P	P	P	P	P	P
Atp6v0a1	ATPase, H+ transporting, lysosomal V0 subun	-	22	1/3	20	98	A	A	A	P	A	A	A	P	P	P	P	P	P
Lefty1	left right determination factor 1	-	23		21	85	A	A	A	A	A	A	A	P	P	P	A	A	P
Ank1	ankyrin 1, erythroid	↑	24	1/4	22	57	A	A	A	A	A	A	A	P	P	P	A	A	P
Rabgef1	RAB guanine nucleotide exchange factor (GEF	-	25	2/3		95	P	P	P	P	P	P	P	P	P	P	P	P	P
Abcb6	ATP-binding cassette, sub-family B (MDR/TA	-	26		23	93	P	P	P	P	P	P	P	P	P	P	P	P	P
30043G21F	RIKEN cDNA C530043G21 gene	-	27	1/2	24	99	P	P	P	P	P	P	P	P	P	P	P	P	P
Abcb10	ATP-binding cassette, sub-family B (MDR/TA	↑	28	3/3		85	P	P	P	P	P	P	P	P	P	P	P	P	P
Ugt8	UDP-glucuronosyltransferase 8	→	29	1/2	25	95	A	A	A	A	A	A	A	P	P	P	P	A	A
Spn1	spectrin beta 1	-	30		26	92	P	A	P	A	A	A	P	P	P	P	A	M	P
Dapk2	death-associated kinase 2	-	31		27	99	A	A	A	A	A	A	P	P	P	P	A	P	P
—	Transcribed locus	-	32		28	?	A	A	A	A	A	A	A	P	P	P	A	A	P
Reps2	RALBP1 associated Eps domain containing pr	-	33		29	?	A	A	A	A	A	A	P	P	P	P	A	A	A
Cpx	coproporphyrinogen oxidase	-	34	2/2		100	P	P	P	P	P	P	P	P	P	P	P	P	P
Mbd2	methyl-CpG binding domain protein 2	→	35	1/2	30	99	P	P	P	P	P	P	P	P	P	P	P	P	P
Epb4.2	erythrocyte protein band 4.2	→	36	2/2		78	A	A	A	A	A	A	A	P	P	P	A	A	P
Dnaj4	DnaJ (Hsp40) homolog, subfamily A, member	-	37	1/3	31	95	A	A	A	A	A	A	A	P	P	P	A	A	A
Dennd2c	DENN/MADD domain containing 2C	-	38		32	88	P	M	P	A	A	P	P	P	P	P	A	P	P
Slc6a9	solute carrier family 6 (neurotransmitter transp	-	39	1/2	33	98	A	A	A	A	A	A	A	P	P	P	A	A	A
Pgm211	phosphoglucomutase 2-like 1	-	40	1/3	34	95	P	P	P	P	P	P	M	P	P	P	P	P	A
Nudt22	nudix (nucleoside diphosphate linked moiety X	-	41	1/2	35	86	P	P	P	P	P	P	P	P	P	P	P	P	P
Pkd212	polycystic kidney disease 2-like 2	-	42		36	83	A	P	A	A	P	A	A	P	P	P	A	A	P
Rps6kb2	ribosomal protein S6 kinase, polypeptide 2	→	43		37	?	P	P	P	P	P	P	P	P	P	P	P	P	P
Pvt1	plasmacytoma variant translocation 1	→	44	1/4	38	?	A	A	A	A	P	M	A	P	P	P	P	A	P
Pon2	paraoxonase 2	→	45		39	91	P	P	P	P	P	P	P	P	P	P	P	P	P
Ankr48	ankyrin repeat domain 48□similar to spermato	-	46		40	?	A	A	A	A	A	A	P	P	P	P	A	A	A
Dnaj4	DnaJ (Hsp40) homolog, subfamily A, member	-	47	2/3		95	A	A	A	A	A	A	A	P	P	P	P	P	P
Rspy1	RIKEN cDNA 4930470D19 gene	-	48	1/2	41	99	P	P	P	P	P	P	P	P	P	P	P	P	P
Copz1	coatamer protein complex, subunit zeta 1	→	49		42	100	P	P	P	P	P	P	P	P	P	P	P	P	P
700047117R	RIKEN cDNA 1700047117 gene	-	50		43	?	P	P	P	P	P	P	P	P	P	P	P	A	P

Appendice H

Gènes spécifiques de E1, E2, E3, Mk3																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	
--------------------------------------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Appendice H suite

[illegible]

Appendice I

Gènes spécifiques de E et Mk																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																													
------------------------------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Appendice I suite

[illegible]

Appendice J

Gènes spécifiques de T, T2, T3																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																	
--------------------------------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Appendice J suite

Gene Symbol	Gene Title	ordre des probesets	probesets redondants	ordre des gènes	% homologie Hs	T1_call	T2_call	T3_call	B1_call	B2_call	B3_call	M_call	E1_call	E3_call	Mk1_call	Mk2_call	Mk3_call
Tcf7	transcription factor 7, T-cell specific	1	1/2	1	92	P	P	P	A	A	A	A	A	A	A	A	A
Cd3z	CD3 antigen, zeta polypeptide	2	1/4	2	86	P	P	P	A	A	A	A	A	A	P	A	A
Tcrim	T-cell receptor interacting molecule	3	1/2	3	72	P	P	P	A	A	A	A	A	A	A	A	A
Camk4	calcium/calmodulin-dependent protein kinase 1	4	1/4	4	92	P	P	P	A	A	A	A	A	A	A	A	A
Cd3z	CD3 antigen, zeta polypeptide	5	2/4		85	P	P	P	A	A	A	A	A	A	A	A	A
Camk4	calcium/calmodulin-dependent protein kinase 1	6	2/4		92	P	P	P	A	P	A	A	A	A	A	P	A
Sh2d1a	SH2 domain protein 1A	7		5	92	P	P	P	A	A	A	A	A	A	A	A	A
Cd3d	CD3 antigen, delta polypeptide	8		6	71	P	P	P	A	A	A	P	A	A	A	A	A
Cd3e	CD3 antigen, epsilon polypeptide	9		7	64	P	P	P	A	A	A	A	A		P	A	A
Gata3	GATA binding protein 3	10		8	98	P	P	P	A	A	A		M		A	A	A
Tcf7	transcription factor 7, T-cell specific	11	2/2		92	P	P	P	P	A	P	M			A	P	P
Tcrb-V13	T-cell receptor beta, variable 13	12	1/7	9	52	P	P	P	P	A	A	P	A		A	A	P
Scap1	src family associated phosphoprotein 1	13	1/4	10	64	P	P	P	P	A	P	A			A	A	A
Cd3g	CD3 antigen, gamma polypeptide	14		11	79	P	P	P	A	A	A	P	M		P	A	P
Tcrb-V13	T-cell receptor beta, variable 13	15	2/7		52	P	P	P	P	A	A	M			A	P	P
---	Transcribed locus, moderately similar to NP_79	16		12	?	P	P	P	P	A	A	A	A		A	A	A
---	16 days neonate thymus cDNA, RIKEN full-len	17		13	?	P	P	P	M		A	A	A		A	A	A
Itk	IL2-inducible T-cell kinase	18	1/3	14	96	P	P	P	A	M		A	A	A	A	A	P
Tcrb-V13	T-cell receptor beta, variable 13	19	3/7		52	P	P	P	A	A	M	M			A	P	P
Lime1	Lck interacting transmembrane adaptor 1	20	1/2	15	66	P	P	P	P	P	P	P	P	P	P	P	P
Lime1	Lck interacting transmembrane adaptor 1	21	2/2		66	P	P	P	A	A	A	P	A		A	A	A
Tcrb-V13	T-cell receptor beta, variable 13	22	4/7		52	P	P	P	P	A	A	A	A		P	A	P
Thy1	thymus cell antigen 1, theta	23		16	74	P	P	P	P	A	P	A	A		P	A	A
E430004N04R	RIKEN cDNA E430004N04 gene	24	1/3	17	85	P	P	P	A	A	A	A	A		A	A	A
Sh2d2a	SH2 domain protein 2A	25		18	70	P	P	P	A	P	A	A	A		A	A	A
Cd28	CD28 antigen	26	1/3	19	72	P	P	P	P	A	P	A			A	P	P
Scap1	src family associated phosphoprotein 1	27	2/4		64	P	P	P	P	P	P	A	A		A	A	A
Tcrb-V13	T-cell receptor beta, variable 13	28	5/7		52	P	P	P	P	A	A	A	A		P	A	P
Itk	IL2-inducible T-cell kinase	29		20	96	P	P	P	P	A	A	A	A		A	A	A
Tcr-gamma	T-cell receptor gamma, constant region	30		21	59	P	P	P	A	A	A	A	A		P	A	M
Tcr-alpha	T-cell receptor alpha chain	31	1/7	22	60	P	P	P	A	P	A	A	A		A	A	A
---	---	32		23	?	P	P	P	P	A	P	P	A	M		A	P
Nr4a1	nuclear receptor subfamily 4, group A, member	33		24	94	P	P	P	P	A	A	P	P		A	P	P
Cd6	CD6 antigen	34		25	74	P	P	P	P	A	A	A	A		A	A	A
E1700017B05R	RIKEN cDNA E1700017B05 gene	35		26	?	P	P	P	P	P	P	P	P	P	P	P	P
Cd28	CD28 antigen	36	2/3		72	P	P	P	P	A	A	A	A		A	P	A
Camk4	calcium/calmodulin-dependent protein kinase 1	37	3/4		92	P	P	P	P	A	A	A	A		A	A	A
Tnfrsf7	tumor necrosis factor receptor superfamily, member 7	38		27	73	P	P	P	A	A	A	M			A	P	P
Ctnnb1	catenin, beta like 1	39		28	98	P	P	P	P	P	P	P	P	P	P	P	P
E430004N04R	RIKEN cDNA E430004N04 gene	40	2/3		85	P	P	P	A	A	A	A	A		A	A	A
E430004N04R	RIKEN cDNA E430004N04 gene	41	3/3		85	P	P	P	A	A	A	A	A		A	A	A
Galk1	galactokinase 1	42		29	90	P	P	P	P	P	P	P	P	P	P	P	P
Cd28	CD28 antigen	43	3/3	30	72	P	P	P	P	A	A	A	A		A	P	A
E2010100O12R	RIKEN cDNA E2010100O12 gene	44		31	100	P	P	P	P	P	P	P	P	P	P	P	P
Pomc1	pro-opiomelanocortin-alpha	45		32	83	P	A	A	A	A	A	A	A		A	A	A
---	---	46		33	?	P	P	P	A	A	A	A	A		A	A	A
Mphosph9	M-phase phosphoprotein 9	47	1/4	34	82	P	P	P	P	P	P	P	P	P	P	P	P
E1500009L16R	RIKEN cDNA E1500009L16 gene	48		35	?	P	P	P	P	A	P	A	A		A	P	A
---	---	49		36	?	P	P	P	P	A	A	A	A		A	A	A
Mphosph9	M-phase phosphoprotein 9	50	2/4	37	82	P	P	P	P	P	P	P	P	P	P	P	P

Appendice K

Gènes spécifiques de T1, T2																		
F1	T1	T1-C	T2	T2-C	T3	T3-C	B1	B2	B3	M	E1	E2	E3	Mk1	Mk2	Mk3	Gene Symbol	
1450545 a at	6,1	-1,1	5,7	-1,5	-1,7	-8,9	-1,7	-1,4	5,4	-1,9	-1,9	-1,7	-1,7	-1,6	-1,9	-1,9	Dntt	T
1427726 at	6,0	0,0	6,6	0,5	-1,1	-7,2	-1,4	-1,4	-1,4	-1,6	-1,3	-1,3	-1,4	-1,1	0,7	-1,4	H2-T3	T
1427655 a at	4,6	-1,4	6,4	0,4	-1,5	-7,5	-1,4	-1,5	-1,4	-1,3	-0,4	-1,5	-1,1	-0,9	-1,2	1,3	630038E17Rik	T
1436115 at	6,8	7,7	2,9	3,8	-1,2	-0,4	-0,8	0,7	-1,3	-1,1	-1,0	-1,1	-0,9	-1,0	-1,0	-1,0	Gm266	
1449937 at	4,4	-1,2	5,5	-0,1	-1,9	-7,5	3,1	1,8	-0,1	-1,0	-1,8	-2,0	-2,3	-1,4	-2,2	-2,1	Pp1lr	
1450521 a at	2,7	5,3	4,3	6,9	-3,2	-0,6	0,2	-1,3	-3,4	-3,8	1,9	1,0	1,5	2,8	-3,5	0,9	Tcrg	T
1449938 at	4,7	-0,6	5,2	0,0	-1,7	-7,0	3,3	2,1	0,0	-1,0	-1,8	-2,4	-2,4	-1,5	-2,3	-2,2	Pp1lr	
1451673 at	4,4	-3,5	6,3	-1,5	-1,1	-9,0	-0,6	-0,7	-1,2	-1,1	-1,0	-1,2	-1,2	-0,7	-1,1	-0,8	Cd8a	T
1449757 x at	5,2	-1,7	4,8	-2,1	-1,7	-8,6	-1,3	-1,4	4,9	-1,8	-1,6	-1,3	-1,5	-1,4	-1,6	-1,4	Dntt	T
1427802 a at	4,7	5,6	6,2	7,1	-1,1	-0,1	-0,1	-0,9	-1,0	-1,1	-1,2	-1,1	-1,0	-1,2	-1,2	-1,0	Tcrg	T
1452548 x at	5,4	0,2	6,0	0,9	-0,8	-6,0	-1,2	-1,5	-1,5	-1,3	-1,2	-0,9	-1,3	-1,0	0,3	-1,1	H2-T3	T
1450680 at	4,7	-2,6	3,9	-3,5	-2,1	-9,5	-1,1	3,6	5,6	-2,2	-1,9	-2,1	-2,1	-1,9	-2,3	-2,0	Rag1	T
1419028 at	3,2	-2,1	4,6	-0,7	-2,3	-7,6	-1,1	-0,4	2,0	-2,2	-1,5	-1,1	-1,8	0,0	0,6	0,0	Arpp21	T
1422189 x at	2,6	5,4	4,4	7,1	-2,6	0,1	-0,1	-1,4	-3,4	-3,4	1,3	0,9	1,8	2,7	-3,6	0,8	Tcrg-V4	T
1425822 a at	4,8	4,7	4,3	4,2	-1,7	-1,8	-0,1	-0,5	-1,2	-1,5	-1,1	-1,4	-1,9	-1,5	0,4	1,4	Dtx1	T
1450924 at	4,8	3,6	5,2	4,0	-1,0	-2,2	-0,9	-1,0	-0,8	-0,9	-1,0	-0,8	-1,0	-0,9	-0,8	-0,8	Hdgfrp3	
1439412 at	2,7	-2,2	5,6	0,7	-1,3	-6,3	-0,9	-0,7	0,7	-1,4	-1,2	-1,2	-1,1	-0,4	-0,6	-0,2	Arpp21	T
1444078 at	3,7	-3,2	5,8	-1,0	-0,9	-7,7	-1,1	-1,1	-0,6	-0,6	-0,9	-1,0	-0,8	-0,6	-1,0	-0,9	Cd8a	T
1442121 at	3,0	-0,7	5,2	1,5	-1,5	-5,2	-1,5	-0,6	1,1	-1,1	-1,4	-1,4	-0,8	-0,8	0,0	-0,2	Arpp21	T
1457656 s at	4,9	-1,0	4,9	-1,0	-0,9	-6,9	-0,6	-0,8	-0,5	-1,0	-1,1	-1,1	-1,3	-1,0	-0,9	-0,7	230085N15Rik	C
1421760 at	4,2	0,6	5,4	1,8	-0,9	-4,5	-0,9	-1,0	-0,9	-1,1	-1,1	-0,9	-0,6	-0,5	-1,0	-0,7	Ptcr	
1415972 at	2,7	0,2	2,9	0,5	-2,9	-5,4	0,6	1,3	1,7	-0,5	-1,5	-1,6	-2,3	-1,0	-0,2	0,8	Marcks	T?
1456700 x at	2,6	0,3	2,8	0,5	-3,0	-5,3	0,5	1,3	1,9	-0,1	-1,2	-1,3	-2,7	-1,3	-0,4	0,7	Marcks	T?
1420692 at	4,9	5,8	4,6	5,5	-0,9	0,0	-0,5	-0,2	0,2	-1,2	-1,2	-1,2	-1,0	-1,0	-1,4	-1,1	Il2ra	T
1435977 at	4,3	3,7	5,5	4,9	-0,6	-1,2	-1,1	-0,9	-1,0	-1,1	-1,1	-1,0	-0,9	-1,1	-0,1	-1,0	Hdgfrp3	
1425335 at	3,1	-4,3	5,6	-1,8	-0,8	-8,3	-0,7	-0,8	-1,0	-0,7	-0,6	-0,8	-0,9	-0,6	-1,0	-0,8	Cd8a	T
1443772 at	4,6	2,5	4,0	1,8	-1,3	-3,4	2,1	0,6	0,1	-1,6	-1,7	-1,4	-1,4	-1,4	-1,5	-1,2	810422M04Rik	2
1452679 at	3,6	-0,7	5,1	0,8	-1,0	-5,3	-0,6	0,6	0,4	-1,4	-1,1	-1,1	-1,1	-1,1	-1,2	-1,2	Tubb2b	
1439897 at	4,1	-2,7	5,5	-1,3	-0,5	-7,3	-0,7	-1,0	-1,0	-0,9	-1,0	-0,8	-1,0	-1,1	-0,8	-0,5	Nebl	
1418734 at	2,5	-4,1	6,1	-0,5	-0,5	-7,1	-0,7	-1,0	-0,8	-0,7	-0,9	-0,9	-0,7	-0,7	-0,9	-0,8	/// H2-Q1 ///	T
1454877 at	3,7	2,5	3,8	2,6	-1,5	-2,7	-0,1	2,8	2,4	-1,8	-1,6	-1,5	-1,8	-1,2	-1,7	-1,6	Sertad4	
1452209 at	3,4	2,1	3,6	2,3	-1,8	-3,0	-1,8	0,3	0,7	-0,3	-1,9	-1,2	-2,1	-0,1	0,3	0,8	Pkp4	
1421965 s at	3,3	2,6	4,2	3,5	-1,5	-2,1	-1,2	-1,0	-1,2	-1,3	-0,6	1,0	1,0	-0,7	-1,2	-0,8	Notch3	T
1448943 at	3,8	1,1	3,9	1,2	-1,3	-4,0	-1,7	-2,1	-2,0	3,4	-0,8	-0,9	-1,3	-1,2	-0,3	0,4	Nrp1	T
1442489 at	3,2	4,4	4,8	5,9	-0,9	0,2	-0,3	-0,5	-0,4	-1,0	-0,7	-1,5	-1,4	-0,6	0,3	-1,0	D1Erd564e	
1428420 a at	3,5	4,9	3,6	5,0	-1,5	-0,1	-1,3	-1,5	-1,3	-1,0	-0,8	-1,2	-1,1	-0,8	3,5	0,0	200009I06Rik	
1418084 at	3,6	1,5	3,9	1,7	-1,2	-3,4	-1,5	-1,4	-1,8	2,9	-1,2	-0,5	-1,4	-1,2	-0,6	0,3	Nrp1	T
1424567 at	3,1	1,3	3,9	2,2	-1,4	-3,2	-1,7	-1,7	-1,1	4,0	-1,1	-1,4	-1,6	0,4	-1,0	-0,4	Tspan2	
1440842 at	4,2	-1,4	4,6	-1,0	-0,5	-6,2	0,1	-0,8	-0,1	-1,3	-1,3	-0,9	-1,0	-1,5	-0,7	-0,8	230085N15Rik	C
1459219 at	2,6	3,1	4,8	5,3	-0,8	-0,4	-0,7	-0,5	-0,8	0,1	-1,0	-1,0	-1,0	-0,8	-0,3	-0,7	Rapgef2	
1435848 at	2,9	2,3	4,8	4,2	-0,8	-1,4	-1,1	-0,6	-0,6	-0,7	-0,8	-0,6	-0,6	-0,3	-0,9	-0,4	---	
1456028 x at	2,3	0,3	2,5	0,5	-2,4	-4,4	0,5	1,0	1,2	0,1	-1,4	-1,1	-2,2	-0,9	-0,2	0,6	Marcks	T?
1444620 at	2,5	0,1	4,3	1,9	-1,1	-3,5	-0,1	-0,6	0,0	-0,9	-0,1	-1,6	-0,7	-0,8	-0,4	-0,5	Tcf12	
1447260 at	3,6	-0,6	4,0	-0,2	-0,9	-5,1	-0,6	-0,6	-0,2	-1,1	-1,0	-1,0	-1,2	-0,9	0,8	-0,9	H2-Q1	
1449682 s at	3,0	-0,4	4,6	1,2	-0,7	-4,1	-0,6	0,0	-0,3	-1,1	-1,0	-0,5	-0,5	-1,0	-1,1	-0,7	410129E14Rik	2
1423252 at	3,5	3,4	4,0	3,9	-0,9	-1,0	-0,7	-0,5	-0,8	-0,6	-0,9	-0,4	-0,9	-1,0	-0,3	-0,6	Hdgfrp3	
1419480 at	2,5	0,8	2,5	0,9	-2,2	-3,8	1,6	2,4	3,0	2,0	-2,5	-4,1	-4,1	-1,5	-0,2	0,6	Sell	T
1459461 at	3,1	4,0	4,5	5,4	-0,7	0,2	-0,4	-0,3	-0,5	-1,0	-0,6	-1,3	-1,3	-1,2	-0,5	0,0	---	
1436398 at	3,9	1,4	5,0	2,5	-0,3	-2,9	1,9	2,1	2,1	-2,0	-1,9	-1,9	-2,2	-2,2	-2,3	-2,2	Lef1	T
1451418 a at	3,4	3,4	3,7	3,7	-1,0	-1,0	-0,2	-1,3	2,6	-0,8	-1,2	-1,2	-1,0	-1,1	-1,0	-1,0	Spsb4	

Appendice K suite

Gene Symbol	Gene Title	ordre des probesets	probesets redondants	ordre des gènes	% homologie Hs	T1_call	T2_call	T3_call	B1_call	B2_call	B3_call	M_call	E1_call	E2_call	E3_call	Mk1_call	Mk2_call	Mk3_call
Dntt	deoxynucleotidyltransferase, terminal (TDT)	1	1/2	1	83	P	P	A	A	M	P	A	A	A	A	A	A	A
H2-T3	histocompatibility 2, T region locus 3	2	1/2	2	?	P	P	A	A	A	A	A	A	A	A	A	P	A
A630038E17Ri	RIKEN cDNA A630038E17 gene	3		3	?	P	P	A	M	A	A	A	P	A	A	A	P	P
Gm266	gene model 266, (NCBI)	4		4	47	P	P	A	P	P	A	A	A	A	P	A	P	M
Pp11r	placental protein 11 related	5	1/2	5	89	P	P	A	P	P	P	P	A	A	A	A	A	A
Tcrg	T-cell receptor gamma	6	1/2	6	?	P	P	P	P	P	A	A	P	P	P	P	A	P
Pp11r	placental protein 11 related	7	2/2		89	P	P	P	P	P	P	P	A	A	A	A	P	A
Cd8a	CD8 antigen, alpha chain	8	1/3	7	51	P	P	A	P	A	A	M	P	A	A	A	P	A
Dntt	deoxynucleotidyltransferase, terminal	9	2/2		83	P	P	A	A	A	P	A	A	A	A	A	A	A
Tcrg	T-cell receptor gamma	10	2/2		?	P	P	A	A	A	A	A	A	A	A	A	A	A
H2-T3	histocompatibility 2, T region locus 3	11	2/2		?	P	P	P	A	A	A	A	A	A	A	A	P	A
Rag1	recombination activating gene 1	12		8	90	P	P	A	P	P	P	A	A	A	A	A	A	A
Arpp21	cyclic AMP-regulated phosphoprotein, 21	13	1/3	9	90	P	P	P	P	P	P	A	P	P	P	P	P	P
Tcrg-V4	T-cell receptor gamma, variable 6	14		10	?	P	P	P	P	P	A	P	P	P	P	P	A	P
Dtx1	deltex 1 homolog (Drosophila)	15		11	99	P	P	A	P	P	P	P	A	A	M	P	P	P
Hdgfrp3	hepatoma-derived growth factor, related protein	16	1/3	12	99	P	P	A	A	A	A	A	A	A	A	A	A	A
Arpp21	cyclic AMP-regulated phosphoprotein, 21	17	2/3		90	P	P	A	P	P	P	A	M	A	A	P	P	P
Cd8a	CD8 antigen, alpha chain	18	2/3		51	P	P	A	A	A	P	P	P	A	A	A	A	A
Arpp21	cyclic AMP-regulated phosphoprotein, 21	19	3/3		90	P	P	A	P	P	P	A	A	A	A	P	P	P
C230085N15Ri	RIKEN cDNA C230085N15 gene	20		13	?	P	P	A	P	P	P	A	A	A	A	A	P	P
Ptcr	pre T-cell antigen receptor alpha	21		14	75	P	P	A	A	A	A	A	A	A	A	A	A	A
Marcks	myristoylated alanine rich protein kinase C subst	22	1/3	15	?	P	P	A	P	P	P	P	P	P	P	P	P	P
Marcks	Myristoylated alanine rich protein kinase C subst	23	2/3		?	P	P	A	P	P	P	P	P	P	P	P	P	P
Il2ra	interleukin 2 receptor, alpha chain	24		16	60	P	P	A	P	P	P	P	P	P	P	P	A	A
Hdgfrp3	hepatoma-derived growth factor, related protein 3	25	2/3		99	P	P	A	A	A	A	A	A	A	A	A	M	A
Cd8a	CD8 antigen, alpha chain	26	3/3		51	P	P	A	A	A	A	A	A	A	A	A	A	A
2810422M04Ri	RIKEN cDNA 2810422M04 gene	27		17	?	P	P	A	P	P	P	A	A	A	A	A	A	P
Tubb2b	tubulin, beta 2b	28		18	100	P	P	P	P	P	P	A	A	A	A	A	A	A
Nebi	nebulin	29		19	?	P	P	A	M	A	M	A	A	A	A	A	A	P
/// H2-Q1 ///	histocompatibility 2, K1, K region /// histocompa	30		20	?	P	P	A	A	A	A	A	A	A	A	A	A	A
Sertad4	SERTA domain containing 4	31		21	84	P	P	A	P	P	P	A	A	A	A	A	A	A
Pkp4	plakophilin 4	32		22	98	P	P	P	P	P	P	P	P	P	P	P	P	P
Notch3	Notch gene homolog 3 (Drosophila)	33		23	?	P	P	A	P	A	A	P	P	P	P	P	A	A
Nrp1	neuropilin 1	34	1/2	24	?	P	P	A	A	A	A	P	P	P	P	A	P	P
D1Erd564e	DNA segment, Chr 1, ERATO Doi 564, expresse	35		25	?	P	P	P	P	P	P	P	P	P	P	P	P	P
1200009I06Ri	RIKEN cDNA 1200009I06 gene	36		26	80	P	P	A	A	A	A	A	P	A	A	P	P	A
Nrp1	neuropilin 1	37	2/2		?	P	P	A	A	A	A	P	A	P	A	A	P	P
Tspan2	tetraspanin 2	38		27	95	P	P	A	P	P	A	P	P	A	A	P	P	P
C230085N15Ri	RIKEN cDNA C230085N15 gene	39		28	?	P	P	A	P	P	P	M	A	A	A	A	P	A
Rapgef2	Rap guanine nucleotide exchange factor (GEF) 2	40		29	97	P	P	A	A	P	A	P	A	A	A	A	P	P
--	--	41		30	93	P	P	A	A	A	A	A	A	A	A	A	A	A
Marcks	Myristoylated alanine rich protein kinase C subst	42	3/3		?	P	P	A	P	P	P	P	P	P	A	P	P	P
Tcf12	RIKEN cDNA E430034C17 gene	43		31	98	P	P	P	P	P	P	P	P	P	P	P	P	P
H2-Q1	Histocompatibility 2, Q region locus 1	44		32	68	P	P	P	P	P	P	A	P	A	A	M	P	P
2410129E14Ri	RIKEN cDNA 2410129E14 gene	45		33	?	P	P	P	P	P	P	A	A	M	A	A	A	A
Hdgfrp3	hepatoma-derived growth factor, related protein	46	3/3		99	P	P	A	A	A	A	A	A	A	A	A	P	A
Sell	selectin, lymphocyte	47		34	80	P	P	P	P	P	P	P	P	A	A	P	P	P
--	--	48		35	?	P	P	P	P	P	P	A	P	A	A	A	P	P
Lef1	Lymphoid enhancer binding factor 1	49		36	?	P	P	P	P	P	P	A	A	A	A	A	A	A
Spsb4	splA/ryanodine receptor domain and SOCS box c	50		37	?	P	P	A	P	A	P	A	A	A	A	A	A	A

Appendice L

Gènes spécifiques de B1, B2, B3																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																													
---------------------------------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Appendice L suite

Gene Symbol	Gene Title	ordre des probés	probés redondants	ordre gfs gènes	% homologie Hs	T1_call	T2_call	T3_call	B1_call	B2_call	B3_call	M_call	E1_call	E2_call	E3_call	Mk1_call	Mk2_call	Mk3_call
Fmn2	formin 2	1	1/2	1	78	A	A	A	P	P	P	A	P	A	A	A	A	A
Vpreb3	pre-B lymphocyte gene 3	2		2	70	A	A	A	P	P	P	A	P	A	A	A	A	A
Bfsp2	beaded filament structural protein 2, phakinin	3		3	88	A	A	A	P	P	P	A	P	A	A	A	A	A
preb1 /// Vpre	pre-B lymphocyte gene 1 /// pre-B lymphocyte	4	1/2	4	75	A	A	A	P	P	P	A	P	A	A	A	A	P
Enpep	glutamyl aminopeptidase	5		5	83	A	A	A	P	P	P	A	P	A	A	A	A	A
Igll1	immunoglobulin lambda-like polypeptide 1	6		6	60	A	A	A	P	P	P	P	P	A	A	A	A	A
---	---	7		7	?	A	A	A	P	P	P	A	P	A	A	A	A	A
---	---	8		8	?	A	A	A	P	P	P	A	P	A	A	A	A	P
Csrp2	cysteine and glycine-rich protein 2	9		9	100	A	P	P	P	P	P	P	P	P	A	A	A	P
Vpreb1	pre-B lymphocyte gene 1	10	2/2		75	A	A	A	P	P	P	A	P	A	A	M	A	A
Ebf1	Early B-cell factor 1	11	1/8	10	100	A	A	A	P	P	P	A	P	A	A	A	P	P
Bst1	bone marrow stromal cell antigen 1	12	1/2	11	77	A	P	A	P	P	P	A	M	A	A	A	A	A
Arntl2	aryl hydrocarbon receptor nuclear translocator	13		12	?	A	A	A	P	P	P	A	P	A	A	A	A	A
Tmem108	transmembrane protein 108	14		13	83	P	A	P	P	P	P	P	P	A	A	A	A	A
Gfra2	glial cell line derived neurotrophic factor family	15	1/3	14	96	A	A	A	P	P	P	A	P	A	A	A	P	P
Ndg2	Nur77 downstream gene 2	16	1/2	15	87	P	P	P	P	P	P	P	P	P	P	P	P	P
Ebf1	early B-cell factor 1	17	2/8		100	A	A	A	P	P	P	A	P	A	A	A	P	P
Serpinf1	serine (or cysteine) proteinase inhibitor, clade	18	1/2	16	89	A	A	A	P	P	P	A	P	A	A	A	A	A
Gfra2	glial cell line derived neurotrophic factor family	19		17	96	A	A	A	P	P	P	P	M	A	A	A	P	P
Ucp2	uncoupling protein 2 (mitochondrial, proton carrier)	20	1/3	18	97	P	P	P	P	P	P	P	P	P	P	P	P	P
Pcmt1	protein-L-isoaspartate (D-aspartate) O-methyltransferase	21	1/7	19	99	P	P	M	P	P	P	P	P	P	P	A	P	P
Ebf1	early B-cell factor 1	22	3/8		100	A	A	A	P	P	P	A	A	A	A	A	P	P
AK129018	cDNA sequence AK129018	23		20	?	A	A	A	P	P	P	A	P	A	A	A	A	A
Bst1	bone marrow stromal cell antigen 1	24	2/2		77	A	P	A	P	P	P	A	A	A	A	A	A	A
Lgr5	leucine rich repeat containing G protein coupled receptor	25		21	89	A	A	A	P	P	P	A	P	A	A	A	A	A
Ndg2	Nur77 downstream gene 2	26	2/2		87	P	P	P	P	P	P	P	P	P	P	P	P	P
Camk2d	calcium/calmodulin-dependent protein kinase	27	1/8	22	100	A	A	A	P	P	P	A	P	A	A	A	A	A
Nedd9	neural precursor cell expressed, developmentally downregulated	28	1/4	23	93	A	A	A	P	P	P	A	P	A	A	A	A	A
Ebf1	Early B-cell factor 1	29	4/8		100	A	P	A	P	P	P	A	P	A	A	A	P	P
Elovl6	ELOVL family member 6, elongation of long chain fatty acid	30	1/5	24	98	P	P	P	P	P	P	A	P	A	A	M	P	P
Kctd3	potassium channel tetramerisation domain containing 3	31	1/2	25	98	P	P	P	P	P	P	P	P	P	P	P	P	P
Plekha1	pleckstrin homology domain containing, family 1	32		26	91	P	P	P	P	P	P	P	P	P	P	P	P	P
Elovl6	ELOVL family member 6, elongation of long chain fatty acid	33	2/5		98	P	P	P	P	P	P	A	P	A	A	A	A	A
930430L01R	RIKEN cDNA 930430L01 gene	34		27	?	A	A	A	P	A	A	A	P	A	A	A	A	A
Rapgef5	Rap guanine nucleotide exchange factor (GEF)	35	1/2	28	95	A	A	A	P	P	P	P	P	A	A	A	A	P
Myo1c	myosin 1C	36	1/4	29	98	M	A	P	P	P	P	P	P	P	A	P	P	P
630023C09R	RIKEN cDNA 9630023C09 gene	37		30	?	A	A	A	P	P	P	A	P	A	A	A	P	P
Gpm6a	glycoprotein m6a	38	1/2	31	99	A	A	A	P	P	P	A	P	A	A	A	A	A
Arpc5l	actin related protein 2/3 complex, subunit 5-like	39	1/2	32	98	P	P	P	P	P	P	P	P	P	P	P	P	P
Elovl6	ELOVL family member 6, elongation of long chain fatty acid	40		33	98	P	P	P	P	P	P	P	A	A	A	P	P	M
Gpm6a	glycoprotein m6a	41	2/2		99	A	A	A	P	P	P	A	P	A	A	A	A	A
Rapgef5	Rap guanine nucleotide exchange factor (GEF)	42	2/2		95	P	P	P	P	P	P	P	P	P	A	A	A	P
Ucp2	uncoupling protein 2 (mitochondrial, proton carrier)	43	2/3		97	P	P	P	P	P	P	P	A	P	P	P	P	P
Snx2	sorting nexin 2	44		34	99	P	P	P	P	P	P	P	P	P	P	P	P	P
Ebf1	Early B-cell factor 1	45	5/8		100	A	A	A	P	P	P	A	A	A	A	A	P	P
810063B05R	RIKEN cDNA 1810063B05 gene	46	1/2	35	79	P	P	P	P	P	P	P	P	P	P	P	P	P
Ebf1	Early B-cell factor 1	47	6/8		100	A	A	A	P	P	P	A	A	A	A	A	A	P
Camk2d	calcium/calmodulin-dependent protein kinase	48	2/8		100	P	P	A	P	P	P	A	P	A	A	A	P	P
Camk2d	calcium/calmodulin-dependent protein kinase	49	3/8		100	A	A	A	P	P	P	A	P	A	A	A	A	A
Slc23a1	solute carrier family 23 (nucleobase transport)	50		36	95	A	A	A	P	P	P	A	P	A	A	A	P	A

Appendice M

Gènes spécifiques des Tet B																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																								
-----------------------------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Appendice M suite

Gene Symbol	Gene Title	log (fold change)	SAM d-score	P-value	FDR	ordre des probets	probets redondant	ordre des gènes	% homology Hs	T1 call	T2 call	T3 call	B1 call	B2 call	B3 call	M call	E1 call	E2 call	E3 call	Mk1 call	Mk2 call	Mk3 call
Fgf13	fibroblast growth factor 13	-5.7	-21.9	<0.001	<0.01	1	1/2	1	99													
Sit1	suppression inducing transmembrane adaptor 1	-4.2	-14.7	<0.001	<0.01	2		2	80	P	P	P	P	P	P	P	P	P	P	P	P	P
Atp1b1	ATPase, Na+/K+ transporting, beta 1 polypeptide	-6.3	-14.3	<0.001	<0.01	3	1/4	3	97	P	P	P	P	P	P	P	P	P	P	P	P	P
Jakmip1	janus kinase and microtubule interacting protein 1	-2.8	-12.7	<0.001	<0.01	4	1/4	4	96	P	P	M	P	P	P	P	P	P	P	P	P	P
Lef1	lymphoid enhancer binding factor 1	-5.4	-12.7	<0.001	<0.01	5	1/4	5	98	P	P	P	P	P	P	P	P	P	P	P	P	P
Ikb3	inhibitor of DNA binding 3	-4.8	-12.5	<0.001	<0.01	6		6	97	P	P	P	P	P	P	P	P	P	P	P	P	P
A630043P06	hypothetical protein A630043P06	-4.8	-12.2	<0.001	<0.01	7		7	?	P	P	P	P	P	P	P	P	P	P	P	P	P
Bcat1	branched chain aminotransferase 1, cytosolic	-4.3	-12.1	<0.001	<0.01	8	1/2	8	89	P	P	P	P	P	P	P	P	P	P	P	P	P
Cdon	Cell adhesion molecule-related/down-regulated by on	-1.6	-12.1	<0.001	<0.01	9		9	86	P	P	P	P	P	P	P	P	P	P	P	P	P
Lck	lymphocyte protein tyrosine kinase	-6.2	-12.1	<0.001	<0.01	10	1/4	10	98	P	P	P	P	P	P	P	P	P	P	P	P	M
Sdc1	syndecan 1	-4.5	-11.9	<0.001	<0.01	11	1/4	11	89	P	P	P	P	P	P	P	P	P	P	P	P	P
Slamf6	SLAM family member 6	-4.7	-11.9	<0.001	<0.01	12	1/3	12	46	P	P	P	P	P	P	P	P	P	P	P	M	P
Brms1l	breast cancer metastasis-suppressor 1-like	-0.8	-11.7	<0.001	<0.01	13	1/2	13	100	P	P	P	P	P	P	P	P	P	P	P	P	P
Dpp4	dipeptidylleptidase 4	-1.8	-11.6	<0.001	<0.01	14	1/5	14	89	P	P	P	P	P	P	P	P	P	P	P	P	P
Celsr1	cadherin EGF LAG seven-pass G-type receptor 1	-3.4	-11.4	<0.001	<0.01	15		15	82	P	P	P	P	P	P	P	P	P	P	P	P	P
Ablim1	actin-binding LIM protein 1	-4.1	-10.9	<0.001	<0.01	16	1/2	16	96	P	P	P	P	P	P	P	P	P	P	P	P	P
Lef1	lymphoid enhancer binding factor 1	-4.3	-10.9	<0.001	<0.01	17	2/4		98	P	P	P	P	P	P	P	P	P	P	P	P	P
'00019D03f	RIKEN cDNA 1700019D03f gene	-1.6	-10.7	<0.001	<0.01	18		17	81	P	P	P	P	P	P	P	P	P	P	P	M	P
Rad51c	Rad51 homolog c (S. cerevisiae)	-2.0	-10.7	<0.001	<0.01	19	1/2	18	88	P	P	P	P	P	P	P	P	P	P	P	P	P
Hnrp3a	heterogeneous nuclear ribonucleoprotein A3	-0.7	-10.7	<0.001	<0.01	20	1/3	19	100	P	P	P	P	P	P	P	P	P	P	P	P	P
Rnf157	Ring finger protein 157	-3.9	-10.6	<0.001	<0.01	21	1/3	20	96	P	P	P	P	P	P	P	P	P	P	P	P	P
Slamf6	SLAM family member 6	-5.4	-10.5	<0.001	<0.01	22	2/3		46	P	P	P	P	P	P	P	P	P	P	P	P	P
Rnf157	Ring finger protein 157	-2.4	-10.4	<0.001	<0.01	23	2/3		96	P	P	P	P	P	P	P	P	P	P	P	P	P
Traf3ip2	Traf3 interacting protein 2	-3.5	-10.4	<0.001	<0.01	24		21	87	P	P	P	P	P	P	P	P	P	P	M	P	P
'30410P16f	RIKEN cDNA 6230410P16f gene	-1.2	-10.4	<0.001	<0.01	25	1/2	22	87	P	P	P	P	P	P	P	P	P	P	P	P	P
Jakmip1	janus kinase and microtubule interacting protein 1	-4.6	-10.2	<0.001	<0.01	26	2/4		96	P	P	P	P	P	P	P	P	P	P	P	M	P
Etv5	ets variant gene 5	-4.5	-9.9	<0.001	<0.01	27	1/3	23	96	P	P	P	P	P	P	P	M	P	P	P	P	P
Lck	lymphocyte protein tyrosine kinase	-7.0	-9.8	<0.001	<0.01	28	2/4		98	P	P	P	P	P	P	P	P	P	P	P	P	P
Sdc1	syndecan 1	-3.8	-9.8	<0.001	<0.01	29	2/4		89	P	P	P	P	P	P	P	P	P	P	P	P	P
Slc39a14	solute carrier family 39 (zinc transporter), member 14	-1.5	-9.7	<0.001	<0.01	30	1/2	24	92	P	P	P	P	P	P	P	P	P	P	P	P	P
Rnf157	Ring finger protein 157	-3.3	-9.7	<0.001	<0.01	31	3/3		96	P	P	P	P	P	P	P	P	P	P	P	P	P
Ablim1	actin-binding LIM protein 1	-5.2	-9.6	<0.001	<0.01	32	2/2		96	P	P	P	P	P	P	P	P	P	P	P	P	P
Dhx30	DEAH (Asp-Glu-Ala-His) box polypeptide 30	-0.5	-9.6	<0.001	<0.01	33	1/3	25	99	P	P	P	P	P	P	P	P	P	P	P	P	P
'30077C05f	RIKEN cDNA 9530077C05f gene	-2.3	-9.6	<0.001	<0.01	34	1/3	26	88	P	P	P	P	P	P	P	P	P	P	P	P	P
Rala	v-rat simian leukemia viral oncogene homolog A (ras	-1.6	-9.6	<0.001	<0.01	35	1/2	27	100	P	P	P	P	P	P	P	P	P	P	P	P	P
Cnot1	CCR4-NOT transcription complex, subunit 1	-0.4	-9.5	<0.001	<0.01	36		28	99	P	P	P	P	P	P	P	P	P	P	P	P	P
Epb4.1l2	erythrocyte protein band 4.1 l-like 2	-1.9	-9.5	<0.001	<0.01	37	1/3	29	91	P	P	P	P	P	P	P	P	P	P	P	P	P
Etv3	ets variant gene 3	-2.1	-9.3	<0.001	<0.01	38	1/4	30	62	P	P	P	P	P	P	P	P	P	P	P	P	P
'30009H04f	RIKEN cDNA A030009H04f gene	-2.1	-9.0	<0.001	<0.01	39	1/2	31	71	P	P	P	P	P	P	P	P	P	P	P	P	P
Atp1b1	ATPase, Na+/K+ transporting, beta 1 polypeptide	-3.4	-8.9	<0.001	<0.01	40	2/4		97	P	P	P	P	P	P	P	P	P	P	P	P	P
Plekha2	pleckstrin homology domain containing, family G (w	-2.3	-8.9	<0.001	<0.01	41		32	84	P	P	P	P	P	P	P	P	P	P	P	P	P
Fgf13	fibroblast growth factor 13	-1.4	-8.8	<0.001	<0.01	42	2/2		99	P	P	P	P	P	P	P	P	P	P	P	P	P
Bcat1	branched chain aminotransferase 1, cytosolic	-2.4	-8.8	<0.001	<0.01	43	2/2		89	P	P	P	P	P	P	P	P	P	P	P	P	P
Whsc1	Wolf-Hirschhorn syndrome candidate 1 (human)	-1.5	-8.7	<0.001	<0.01	44	1/2	33	98	P	P	P	P	P	P	P	P	P	P	P	P	P
Hells	helicase, lymphoid specific	-1.4	-8.7	<0.001	<0.01	45		34	?	P	P	P	P	P	P	P	P	P	P	M	P	P
Edaradd	EDAR (ectodysplasin-A receptor)-associated death d	-3.9	-8.7	<0.001	<0.01	46	1/3	35	82	P	P	P	P	P	P	P	P	P	P	P	P	P
'130054H0i	hypothetical protein D130054H0i	-0.9	-8.5	<0.001	<0.01	47		36	54	P	P	P	P	P	P	P	P	P	P	P	P	M
Galnt12	UDP-N-acetyl-alpha-D-galactosamine polypeptide N	-3.9	-8.5	<0.001	<0.01	48		37	89	P	P	P	P	P	P	P	P	P	P	P	P	P
Ints2	integrator complex subunit 2	-0.9	-8.4	<0.001	<0.01	49		38	?	P	P	P	P	P	P	P	P	P	P	P	P	P
Il7r	interleukin 7 receptor	-5.9	-8.4	<0.001	<0.01	50	1/2	39	69	P	P	P	P	P	P	P	P	P	P	P	P	P